

A Simple Model for Global H I Profiles of Galaxies.

I. M. Stewart^{1,2}, S.-L. Blyth² and W. J. G. de Blok^{3,2,4}

¹ Argelander Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany
e-mail: ims@astro.uni-bonn.de

² Astrophysics, Cosmology and Gravity Centre, Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa.

³ Netherlands Institute for Radio Astronomy (ASTRON), Postbus 2, NL-7990 AA Dwingeloo, The Netherlands.

⁴ Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, The Netherlands.

May 15, 2014

ABSTRACT

Context. Current and future blind surveys for H I generate large catalogs of spectral lines for which automated characterisation would be convenient.

Aims. A 6-parameter mathematical model for H I galactic spectral lines is described. The aim of the paper is to show that this model is indeed a useful way to characterise such lines.

Methods. The model is fitted to spectral lines extracted for the 34 spiral galaxies of the recent high-definition THINGS survey. Three scenarios with different instrumental characteristics are compared. Quantities obtained from the model fits, most importantly line width and total flux, are compared with analog quantities measured in more standard, non-parametric ways.

Results. The model is shown to be a good fit to nearly all the THINGS profiles. When extra noise is added to the test spectra, the fits remain consistent; the model-fitting approach is also shown to return superior estimates of linewidth and flux under such conditions.

Key words. Line: profiles – Methods: data analysis – Methods: numerical – Radio lines: galaxies

1. Introduction

Since hydrogen is the most common element in the universe, and makes up most of its mass, the distribution of this element is intimately connected with the dynamics of matter on scales from the cosmic down to the stellar. Although in its ionized or molecular forms hydrogen is problematic to detect, neutral atomic hydrogen (H I) emits (or absorbs, if it is cooler than the background), due to a hyperfine transition, radio waves at a rest frequency of 1420 MHz. The transition rate is very low, which makes the line faint and relatively difficult to detect, but this has two compensating advantages: firstly, clouds of hydrogen up to the galactic scale are usually optically thin, which makes it relatively straightforward to deduce the mass of the cloud from the intensity of the line; and secondly that the spectral line is intrinsically very narrow, making it a useful way to study the motions, whether thermal, internal or bulk, of the emitting cloud, via Doppler shift and broadening.

In recent years there have been a number of blind surveys for H I in the local universe, some complete (e.g. HIPASS, Barnes et al. 2001), others ongoing (ALFALFA Giovanelli et al. 2005; EBHIS Winkel et al. 2010; Kerp et al. 2011; CHILES Fernández et al. 2013). Further, deeper blind searches for H I (LADUMA Holwerda et al. 2012; WALLABY Koribalski & Staveley-Smith 2009; DINGO Meyer 2009; see also Duffy et al. 2012) are planned with the upcoming SKA precursor instruments MeerKAT (Jonas 2009) and ASKAP (Johnston et al. 2008). Such surveys complement those performed at other wavelengths and avoid potential biases from preferential selection of galaxies which are bright at these wavelengths.

The majority of hydrogen in the universe occurs in galaxies. The width of a galaxy's H I spectral profile gives its bulk rotation

speed, and the area under the profile is proportional to its H I mass – both quantities of importance in cosmology.

From simple geometrical considerations it is clear that in any survey of objects uniformly distributed in space, the number frequency of objects will increase as their angular size decreases. Since the detectability of objects takes a sharp downturn as their angular sizes become smaller than the instrument resolution, the most common survey object therefore can be expected to be one which is only just brighter than the survey sensitivity cutoff, and which has an angular size no larger than the beam of the instrument. Since detection of H I sources usually makes use also of spectral information, a galaxy at the limit of detection may be unresolved in any one channel map, but nevertheless kinematically resolved, such that the flux peak moves progressively across channels (see the EBHIS observation of DDO 154 described in section 3.3.3 for an example of this).

The principal task therefore for post-detection processing of blind H I surveys is to measure as accurately as possible the width, area and other properties of low signal-to-noise (S/N) spectral lines of unresolved sources: and not only for spatially-integrated spectral lines, but also for the line at each spatial pixel across the extent of the source. This is the aim of the technique discussed in the present paper.

Fitting of a model to a spatially-integrated spectral profile where the source is well-resolved on the other hand should not be expected to yield better values of galaxy parameters than other methods, because with such sources there are problems determining which spatial pixels to include in the sum – significant numbers of pixels with small but non-zero contributions may be excluded, or pixels containing nothing but a chance spike in noise may be mistakenly included. We make use of the well-resolved THINGS observations here for practical, demonstrative

reasons. It is unavoidable that pixel masking effects cause flux biases in such spectra, although careful treatment of the data can minimize this.

The question of spatial masking or selection is one that arises even if the source is unresolved, because its flux remains distributed across the imaging beam or point spread function (PSF) of the instrument. Any cutoff criterion will miss some contribution in the wings, but a too generous cutoff will include too much noise from off-source directions. Fitting a model which includes a spatial component shaped like the PSF seems like a simple way around this, but is unsatisfactory because the noise in adjacent spatial pixels is also convolved by the PSF, and is therefore not statistically independent. Statistically speaking, it is better to fit to data which is as unprocessed as practical - i.e., to include instrumental response in the model and fit to raw data, rather than to fit a simpler model to data which has been processed (with accompanying muddying of the statistical waters) in order to remove or at least systematize the instrumental response. But an exploration of such matters is beyond the scope of the present paper. Here we content ourselves with construction of global spectra by spatial masking and summing, and make the associated caveats about the reliability of resulting flux measurements.

The plan of the paper is as follows. The advantages of having a model of the H I spectral line is discussed in Sect. 2.1, and the model itself is described in Sect. 2.2. Various technical matters connected with the process of fitting, including the Bayesian methodology, estimation of uncertainties, and a suitable goodness-of-fit measure, are discussed in Sect. 2.4. The tests performed on the model are presented in Sect. 3. H I spectral lines were obtained under a variety of conditions of noise, spectral resolution and baseline, and were fitted by the model.

We posed three questions about the fits. Firstly, how good a fit is the model in the case of optimum frequency resolution and S/N - does it return good values of bulk properties of the galaxies? This question is addressed in Sect. 3.1, where we fit the model to a spatially-integrated spectrum from each of the 34 THINGS galaxies observed by Walter et al. (2008).

Secondly, if we add noise to the spectra, and bin the channels more coarsely, does this affect the fitted parameter values? We address this question in Sect. 3.2.

Thirdly, how does the model perform fitting to a single-dish observation, where it is necessary also to fit a baseline? And can the model return useful kinetic information when the galaxy is at the limit of resolution? In Sect. 3.3, data from the EBHIS survey (Winkel et al. 2010; Kerp et al. 2011) are used to test this.

2. A model of the H I spectral line

2.1. Introduction

The aim of any model of a physical process is to approximate that process with a small, hence manageable, number of adjustable parameters. A good model will be a good match to the data; will have a small number of parameters; and it is also useful if the model is not just empirical but tied in some way to the underlying physics of the object. Since these desirables can be in opposition, a model is usually then also a compromise.

The shape of an H I spectral line represents the distribution and motions of neutral hydrogen within a galaxy. On top of the bulk rotation which prevents gravitational collapse there can be innumerable variations in the pattern of local H I flow from galaxy to galaxy, which will be reflected in differences between the shapes of the respective spectral lines. It might therefore

seem difficult to formulate a simple model of the H I spectral line. However as shown in Sect. 3.1.2, in practice the 6-parameter model presented here works fairly well. Local deviations in flux density between the data and the fitted model don't exceed about 10%, and tend not to affect either the total flux or the linewidth.

In all the H I surveys the authors are aware of, total flux and linewidth have been estimated from the data directly, without fitting a model to the line profile.¹ So why use one? Firstly because it is more systematic: there is no need for either human intervention or ad-hoc prescriptions for the number of channels to consider. Secondly, as is shown in Sect. 3.2, the bias in linewidth measurement is much reduced through model fitting. Thirdly, as shown in Sect. 3.3, the modelling approach very naturally accommodates a modelling of the spectral background or baseline. It's no longer necessary to decide where the line profile 'ends' before fitting the baseline, since both can be considered together. Fourthly, fitting of a model arguably lends itself more easily to automated processing, which becomes an ever more pressing consideration as survey datasets grow in size.

Lastly, a parametrized model opens the door to the use of Bayesian techniques, which are becoming increasingly accepted as useful tools in astronomy. Advantages here fall under three main heads. The first of these is that a Bayesian formulation is the formally correct (and therefore optimum) procedure for estimating the parameters of interest and for incorporating prior knowledge. This method is applied in the present paper. The second head or category is the use of Bayes' theorem to assess the relative suitability of differing models. This is also used in the present paper, to determine the best order of baseline model (Sect. B.6). Note that the same formulation can be used to estimate detection probability directly, although this topic is not explored in depth in the present paper: the Bayesian approach automatically takes into account both prior knowledge of the expected range of line shapes, as well as the total bandwidth and area of the survey. Such an approach to detection is more fundamental and rigorous than relying on either the 5-sigma rule or ad hoc prescriptions for calculating signal-to-noise ratio (S/N). The third advantage of the Bayesian formalism is in the extraction of statistical results from large ensembles of low-quality data via hierarchical modelling (Loredo 2012). This technique is however not used here.

Roberts (1978) discussed the H I spectral line and explained why it often has a double-horned shape with a relatively steep rise and fall. Roberts explored several semi-physical models of the line profile, and in fact his model C is a special case of ours. Roberts was concerned rather to emphasize and explore the connection between galactic linewidth and luminosity now more usually associated with Tully & Fisher (1977), and his profile models don't have the flexibility to fit a wide variety of H I line shapes.

In the second of two papers which presented a comprehensive simulation of gas in the local universe, Obreschkow et al. (2009b) described a profile model (which they apply to both H I and CO spectral lines) which uses 5 parameters. These parameters, labelled by the authors k_1 to k_5 , are purely empirical in themselves, but can be derived from more fundamental properties of the line profile via a set of formulae (equations A2 to A6, appendix A of their paper). This second tier of parameters includes the flux density at profile centre, the maximum flux den-

¹ Saintonge (2007) described a model constructed from Hermite polynomials which is used in the ALFALFA source-detection procedure (Haynes et al. 2011), but these authors do not, so far as we are aware, make further use of the fit parameters.

sity, and the velocity widths at the 50% and 20% levels. In an earlier paper by some of the same authors (Obreschkow et al. 2009a), the second-tier parameters are derived by constructing each profile from an appropriate projection and integration of a comprehensive model of the distribution of H I gas density and velocity as a function of radius from the galaxy centre. This radial model in turn depends on several third-tier fundamental parameters, such as masses and characteristic radii for disk, halo and bulge components. It might be possible to find a reasonably simple way to connect these fundamental parameters with the eventual k values, but these authors did not attempt this themselves. Their intent was to generate the second-tier parameters for a large number of simulated galaxies and make these available in a database. The k -prescription was provided simply as a convenience for the user who wished to construct approximate profiles from these data.

One could make use of the k -model of Obreschkow et al. for fitting to real data, but there are a number of desirable improvements:

1. The k parameters provide no direct physical insight.
2. Many real H I spectral lines exhibit a noticeable asymmetry (see e.g. Richter & Sancisi 1994). The k model does not cater for this.
3. Since the prescribed functions of the k model are discontinuous at sharply-defined velocity bounds, not aligned *a priori* with velocity channel boundaries, integrating the model over a finite channel width is a little fiddly.
4. There is no way to simulate any artefacts arising from the autocorrelation process which generates spectra from radio signals; nor can the k -model simulate applied filters, such as Hanning or Tukey filters.

The model described in the present paper addresses all these issues.

2.2. The model

In order to derive a model we need to arrive at a reasonable approximation to the motion of H I in a generic galaxy. We will begin by breaking the motions into two categories: bulk vs. random.

2.2.1. Bulk motions

We approximate bulk motions firstly by assuming that all gas in a (late-type, therefore gas-rich) galaxy rotates about a common centre and in a common plane. As shown for example in de Blok et al. (2008), this is a reasonable assumption for the majority of spiral galaxies. We don't assume that the density of gas is azimuthally symmetric: departures from same are catered for, albeit in a crude manner, via the asymmetry parameter of our model.

As is well known, for most galaxies with a total H I mass larger than about 10^9 solar masses, the curve of rotation speed as a function of radius from the centre is seen to be remarkably flat outside the core. We assume here that the rotation curve is always exactly flat outside a certain radius. Within the core itself, the velocity usually rises steadily from (nominally) zero at the galactic centre, then turns over smoothly when it reaches the 'flat' value of velocity. In the present model we approximate this rise by a straight line, such as would be observed for example in a rotating solid body; we also assume that the density of gas in the core is uniform. This simplified rotation curve is similar in shape to that labelled 'C' in Fig. 3 of Roberts (1978).

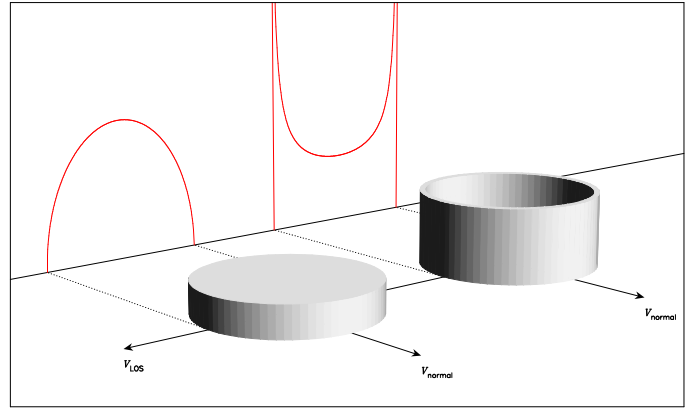


Fig. 1. A schematic showing the model distribution of H I in phase space for an edge-on galaxy. The disk and the ring represent respectively the inner and outer components of the model. The height of each body represents the density of H I as a function of velocity. v_{LOS} is the velocity in the line of sight and v_{normal} is the velocity normal to that, but in the plane of the galaxy. The figures in red represent the H I densities due to the two components, as projected onto the line of sight. Since the densities are displayed as functions of velocity, these red projections in fact give directly the spectral line shapes of the two components. Notes: (i) The ring formally speaking ought to be infinitely thin, but some visible width has been given to it for the sake of easier interpretation. (ii) No attempt has been made to make the height scales of the projection graphs consistent with those of the density figures.

Understanding of the way a profile model represents the bulk gas motions is assisted by mapping the gas motions and densities in phase space - that is, presenting the gas density not as a function of spatial location but of velocity components. The line profile may then be obtained by projecting the phase space gas distribution onto a line directed towards the observer. This is diagrammed in Fig. 1. In phase space, the outer mass of gas with a constant rotation speed appears as a ring of infinitesimal thickness, whereas the 'solid-rotating' inner part appears as a uniform disk.

The ring gives in projection the characteristic double-horned profile so often observed in H I spectral lines:

$$s_{\text{outer}}(v) = \frac{2S}{\pi\Delta v} \rho^{-1} \left(\frac{v - v_{\text{ctr}}}{\Delta v/2} \right) \quad (1)$$

where

$$\rho^2(u) = 1 - u^2 \text{ for } |u| < 1, \\ = 0 \text{ else.}$$

Here S is the total flux from the H I in this section, Δv is the range between maximum and minimum gas velocities in the line of sight, and v_{ctr} is the mean line-of-sight velocity. The disk in projection yields a half ellipse:

$$s_{\text{inner}}(v) = \frac{4S}{\pi\Delta v} \rho \left(\frac{v - v_{\text{ctr}}}{\Delta v/2} \right).$$

These are the two fundamental components of our model. Asymmetry in the line profile is accommodated by multiplying both components by

$$1 + 2\alpha(v - v_{\text{ctr}})/\Delta v$$

To describe the model so far we require 5 parameters: the line centre v_{ctr} ; the so-called intrinsic line width Δv ; the total flux

S ; the fraction f of the gas which is found in the ‘solid-rotating’ part which is associated with the inner part of the galaxy; and the asymmetry parameter α . The model so far is thus represented by

$$s_{\text{intrinsic}}(v) = \frac{2S}{\pi\Delta v} \left[1 + \frac{2\alpha(v - v_{\text{ctr}})}{\Delta v} \right] \times \left[(1-f)\rho^{-1}\left(\frac{v - v_{\text{ctr}}}{\Delta v/2}\right) + 2f\rho\left(\frac{v - v_{\text{ctr}}}{\Delta v/2}\right) \right]. \quad (2)$$

2.2.2. Random motions

Random motions include both thermal motions of the atoms and turbulent motions. The latter may occur on length scales much larger than the atomic, but provided they are much smaller than the scale of bulk, i.e. galaxy-scale motions, we can lump them together with thermal motions. In the present study we assume that the distribution of random velocities is the same throughout any galaxy. Maps of the local linewidth in Walter et al. (2008) indicate that this is a reasonable assumption for the THINGS galaxies. In this approximation, random motions may be treated mathematically as a 3-dimensional convolution of the bulk-motion line profile. Further, if we approximate the distribution of random velocities by a 3D Gaussian, then the 1D projection of this in any direction is also a Gaussian:

$$g(v) = \frac{1}{\Delta v_{\text{rand}} \sqrt{2\pi}} \exp\left(-\frac{v^2}{2[\Delta v_{\text{rand}}]^2}\right). \quad (3)$$

The characteristic width Δv_{rand} is the sixth and final model parameter.

It should be emphasized that representation of random gas motions by a single Gaussian is only an approximation. In reality there may be several co-located components exhibiting different degrees of dispersion (see for example Braun 1997).

The full equation for the profile model is

$$s(v) = s_{\text{intrinsic}}(v) \star g(v) \quad (4)$$

where \star indicates convolution. The full list of 6 model parameters is v_{ctr} , Δv , S , Δv_{rand} , f and α . These have natural ranges, that is, ranges imposed by physical reasonableness, as follows:

- v_{ctr} : constrained in practice by the ends of the spectrum.
- $\Delta v \geq 0$
- $S > 0$
- $\Delta v_{\text{rand}} > 0$
- $0 \leq f \leq 1$
- $-1 \leq \alpha \leq 1$

Note however that, for some of the fits reported in the present paper, up to 6 additional parameters were used for fitting the baseline.

The asymmetry and fraction-solid parameters can be expected to be correlated respectively with the third and fourth moments of the spectral line, as described by Andersen & Bershadsky (2009). These authors make a comprehensive study of the statistics of H I and H II spectral line shapes and their relation to galaxy morphology. We don’t here retrace that analysis, but only observe that calculation of the appropriate moments of our spectral line model is easy and avoids the necessity, mentioned by Andersen & Bershadsky, of *a priori* human selection of a velocity range, when extracting moments from data. Other measures of asymmetry, such as that of Tift & Cocke (1988), are equally easy to calculate from our model.

2.3. Fourier-space formulation

Although the model can be calculated directly using equation 4 and the preceding formalism, there are several advantages to calculating the profile first in Fourier space, then transforming to velocity space. Firstly, the convolution in equation 4 turns into a product; secondly, the singularities in equation 1 are avoided; and thirdly, the process mimics the processing of real signals in an XF-type correlator, and thus allows inclusion of some of the artefacts which result from same.

This Fourier-space model construction was followed in all the calculations described in the present paper. To make it easy for readers to do this themselves, the Fourier transform of equation 4 is given in appendix A.

2.4. Fitting considerations

A frequent use for such a model is to fit it to data. For this purpose one needs an objective function describing the goodness of the fit, and one must choose an algorithm for minimizing this function. These considerations are discussed in Sects. 2.4.1 through 2.4.4.

2.4.1. Bayesian formulation

According to Bayes’ theorem, given a set of measurements of flux density \mathbf{y} , the posterior probability density function of the model parameters $p(\mathbf{q}|\mathbf{y})$, where we use \mathbf{q} as shorthand for the six model parameters, is given by (see e.g. D’Agostini 2003)

$$p(\mathbf{q}|\mathbf{y}) = \frac{1}{E} p(\mathbf{q}) p(\mathbf{y}|\mathbf{q}), \quad (5)$$

where E , known as the evidence, is just a normalizing constant:

$$E = \int d\mathbf{q} p(\mathbf{q}) p(\mathbf{y}|\mathbf{q}). \quad (6)$$

The first function in the integrand is the prior probability distribution which represents our prior knowledge of the parameter values; the second is the likelihood. For N data values y_j which include Gaussian-distributed noise, the likelihood is given by

$$p(\mathbf{y}|\mathbf{q}) = \prod_{j=1}^N \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left\{-\frac{[y_j - s(v_j, \mathbf{q})]^2}{2\sigma_j^2}\right\} \\ = (2\pi)^{-N/2} \exp\left(-\frac{\chi^2}{2}\right) \prod_{j=1}^N \frac{1}{\sigma_j} \quad (7)$$

where $s(v_j, \mathbf{q})$ is the profile model evaluated for velocity channel j , σ_j is the standard deviation of the noise in channel j , and χ^2 has its usual formulation.

A Bayesian fit (that is, a fit procedure which optimizes the Bayesian posterior probability) may tend towards being data-dominated, or it may be prior-dominated. The first case occurs if the data has high signal-to-noise (S/N) and if there has not been much earlier fitting experience. The original THINGS dataset observed with the VLA matches this criterion for us. The EBHIS observations of the THINGS galaxies however generally have much lower S/N values, as do the semi-simulated profiles described in Sect. 3.2. For the latter two cases we therefore thought it appropriate to set some cautious priors, derived from the fits to the VLA THINGS profiles. Some of the model parameters are poorly constrained by the data in these lower-S/N fits and these are thus prior-dominated.

The low-S/N priors are described in detail in appendix B.2.

2.4.2. Fitting algorithms

Three fitting techniques have been used in the present study: the Levenberg-Marquardt (LM) method (Press et al. 1992, chapter 15.5), simplex optimization (Nelder & Mead 1965, see also Press et al. 1992 chapter 10.4) and Markov-chain Monte Carlo (MCMC), specifically using the Metropolis algorithm (Metropolis et al. 1953, see also Bhanot 1988 for other useful references). For fitting to the THINGS profiles, both from the original VLA observations (Sect. 3.1) and as observed as part of EBHIS (Sect. 3.3), the LM procedure was used to obtain initial values of the approximate centre and width of the posterior, then the exact form of the posterior was explored via MCMC. For the semi-artificial profiles described in Sect. 3.2, the simplex procedure alone was used.

The simplex algorithm is robust but slow. In practice it was ‘fast enough’, requiring only about 10 to 20 seconds on a standard laptop to fit a 6-parameter model to 450 data points.

Technical details of the MCMC are given in appendix B.4.

2.4.3. Uncertainties

The parameter uncertainties quoted in Table 1 are the square roots of the diagonal elements of a covariance matrix estimated from the ensemble of MCMC points. For the LM and simplex fits, the covariance matrix was obtained where necessary by inverting the matrix of second derivatives (known as the Hessian) of the posterior with respect to the parameters. As can be seen in the figures in appendix C, the MCMC values are often about 50% larger than the LM ones. The reason for this is not known.

All flavours of uncertainty scale with the uncertainties in the original values of flux density. The calculation of these is described in appendix B.5.

2.4.4. Goodness of fit

The value of χ^2 is commonly used to assess how well a model fits the data. In effect, reduced χ^2 is a measure of the ratio between the data-minus-model residuals and the amplitude of the measurement noise. Noise plays the vital role in this because it affects the probability that the observed residuals would occur by chance with a perfect fit. In the present case however this is not quite what we want. We expect from the start that the model will not be a perfect fit to the data, so in a sense this question is already decided in the negative: what we want is rather to measure the size of the imperfections. Noise plays no role in this, provided only that it is not so large as to swamp the deviations; hence the bare value of χ^2 will not return the information we want.

The line shape of NGC 925 in Fig. 2 provides a good example of the issue. Clearly seen are many local ‘bumps’ or ‘wiggles’ in the H I distribution which the model is not able to track. It is the relative size of these bumps and dips compared to the average height of the profile which it would be most useful to know. For example, from Fig. 2 one can clearly see that the model is a cleaner fit to NGC 3184 than to NGC 925: what we want is some formula to quantify this difference.

The formula we devised to estimate the wiggle fraction J is

$$J^2 = \frac{1}{S^3} \left(\frac{\Delta v}{\Delta v_{\text{chan}}} \right)^2 \sum_j y_j \left[(y_j - s_j)^2 - \sigma_j^2 \right]. \quad (8)$$

where S and Δv are respectively the total flux and line width model parameters as described in Sect. 2.2, and Δv_{chan} is the

width of the spectral channels. The reasoning behind this formula is as follows. Firstly, we want the square of the residual in each channel. We then want to subtract the noise from this in quadrature. The resulting term should be weighted by the flux density y_j of the line profile. This is normalized by dividing by the sum S of the model profile values. The result so far is the square of the average residual, within the line profile, due solely to model/data mismatch. Finally the square root is taken and that result divided by the average flux density of the model, which is

$$\langle s \rangle = \frac{\Delta v_{\text{chan}}}{\Delta v} S.$$

The result is the average fractional residual due to ‘wiggles’. This value is given in Col. 8 of Table 1.

Local fluctuations in H I can be visually deceptive. A good example is the global spectrum of NGC 4214, which is shown together with a model fit in Fig. 2. At first inspection it is not clear why the model has not better fitted the seemingly regular double-horned profile. However, the linewidth of this galaxy is, at 56 km s⁻¹, relatively small: less than 5 times the ‘sigma’ width of the turbulent broadening, which here is 12 km s⁻¹. Further consultation of Table 1 shows that the fraction of solid rotation fitted (which tends to reduce the depth of the valley between the horns) is small - consistent with zero. In fact the fitting procedure has chosen the sharpest and deepest possible double-horned profile which, after smoothing by the turbulent-broadening Gaussian, is consistent with the line slopes. With this amount of smoothing it is not possible for the model to follow deviations over velocity scales less than 10 km s⁻¹, as occurs with NGC 4214. These fluctuations only fool the eye into assuming a regularity which isn’t in fact there, because NGC 4214 is too narrow to make obvious the actually random nature of its wiggles; in contrast to the spectrum of NGC 925, for example (shown in the same figure).

It is worth observing too that, although the wiggles in NGC 4214 do rather offend the eye, the J factor for this galaxy is only 2.3%, well below the average seen in Fig. 3.

3. Tests of the model

3.1. Spectra from the original THINGS observations

3.1.1. Introduction

Walter et al. (2008) used the VLA to observe the H I distribution in 34 nearby late-type galaxies. Spatial resolution, frequency resolution and S/N were all relatively high, certainly when compared to the most common objects detected in blind H I surveys. The set of observations is known as THINGS (The H I Nearby Galaxy Survey). We fitted our profile model to a spectrum extracted from THINGS observations of each of the 34 targets.

Walter et al. (2008) provide these spectra in the online version of their paper. According to their description, the global spectral profile for each galaxy is obtained by adding together a subset of pixels for each channel of the data cube. Pixels where there was no measurable emission from H I were not included in the sum. This has the effect of making the noise in a channel proportional to the square root of the number of unmasked pixels. The mask cubes are not available on the THINGS web site but were kindly provided on request by F. Walter. This allowed us to estimate the noise per channel for each global profile. The exact procedure for doing so is described in appendix B.5.

Some of the THINGS galaxies have a velocity range which straddles zero. For some of these, interference from Milky-Way H I is evident. A few channels where this effect was obvious have

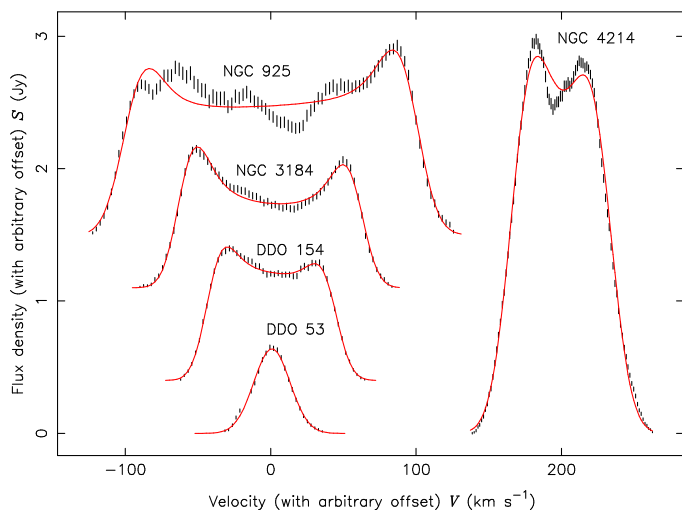


Fig. 2. Fitted profiles (red curves) compared to raw data (black error bars) for 5 of the THINGS galaxies. Widths and heights of the profiles have not been altered but arbitrary offsets to both flux density and velocity have been added for clarity of plotting.

been omitted when fitting to the spectra for galaxies DDO 53, NGC 2976, NGC 3077 and NGC 6946. Note that the total flux values in Col. 4 of Table 1 are the values of the respective parameter of the fitted model: thus these represent an interpolation over missing channels for the four galaxies mentioned. One can however also obtain the total flux under the fitted profile by adding together the samples of the profile model obtained at each channel. For Fig. 4, in which the flux under the data spectrum is compared to the model, the relevant channels have been omitted from these sums for both the data and the model for the four galaxies mentioned. The percentage differences in model flux from the values in the Table are respectively 7, 13, 4 and 3.

The question discussed in the remainder of Sect. 3.1 is: how good a fit is the model in the case of optimum frequency resolution and S/N - does it return good values of bulk properties of the galaxies?

3.1.2. Graphical and tabular display of fit results

Five examples of profiles fitted to THINGS spectra are shown in Fig. 2. These are the five galaxies chosen as ‘simulation inputs’ in Sect. 3.2. Figures showing all 34 individual fit results are given in appendix C.

The results of the fits are given in Table 1. Shown in Cols. 2 to 7 are the mean values with uncertainties for each of the six model parameters. Also given in Col. 8 is the ‘wobble fraction’ J calculated from equation 8. A low value indicates a good match between model and data. Note that it is possible for J^2 to be negative. In this case, formally speaking, the root is imaginary, as listed in the table. This has no physical meaning, it just indicates that any local deviations in H I from the model are insignificant compared to the measurement noise.

A histogram showing the distribution of wobble fraction is given in Fig. 3.

Column 10 of the table gives the velocity separation between points on the fitted profile where the flux density decreases to 20% of its maximum value. The remaining columns, 9 and 11, are described in the following section.

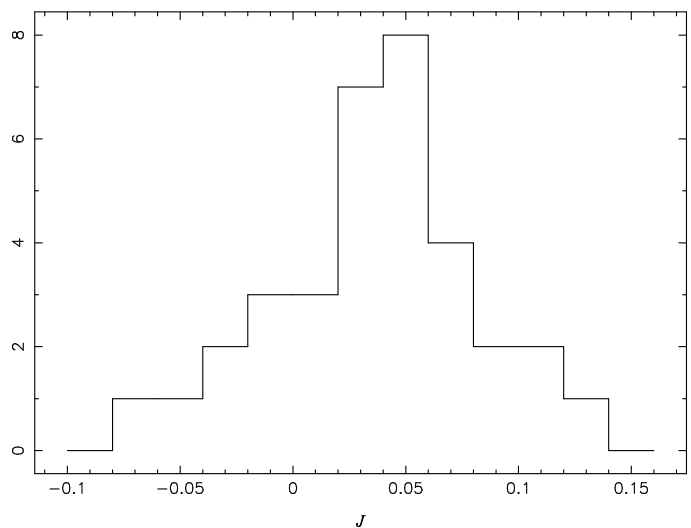


Fig. 3. A histogram showing the distribution of ‘wobble fraction’ J among the THINGS galaxies, as specified by equation 8. This quantity indicates the approximate extent of local fluctuations of H I density away from the fitted model. Values shown here as less than zero are technically imaginary (see the respective values in Col. 8 of Table 1), since they are square roots of differences in quadrature which turned out to be negative. Such values simply indicate that fluctuations away from the model fit are, for that galaxy, dominated by measurement noise.

3.1.3. Numerical comparisons between model and data

The reason for doing these fits is to see how well the model matches a variety of real spectra. One can gain a qualitative impression from looking at the profiles but it is more informative to perform some numerical comparisons between the fitted model parameters and equivalent values from other sources. Some examples are shown in Figs. 4 to 7, which are described individually below.

Figure 4 compares the total flux S_{fit} from summing valid values of the fitted flux density to a similar sum S_{data} over the raw data values. What is displayed in the figure is the fractional difference between these ‘fit’ and ‘data’ flux values for each galaxy.

For about half of the galaxies, the difference between fluxes S_{data} and S_{fit} as displayed in Fig. 4 is consistent with zero. Almost all of the galaxies have flux differences less than about 1%. Where there is a perceptible difference, the total flux values for the data (S_{data}) are always larger than for the fitted profile.

No detailed explanation for this flux anomaly is known at present, but any explanation ought to start with the observation that the model is nonlinear in some of its parameters, and that it doesn’t have support across the whole spectrum. Worth particular attention are the line wings, which in the model become closer to Gaussian in shape the further away from line centre one goes. This Gaussian component is intended to model the distribution of random gas velocities. However, it is known that the true distribution of velocities is better described by a sum of at least 2 Gaussians (Ianjamasimanana et al. 2012). It is not hard to see that a Gaussian which is the best fit to the steep line edges might nevertheless miss flux present in broad, non-Gaussian wings for example.

As a toy model, with a non-linear parameter and limited support, to demonstrate how a pedestal can lead to flux underestimation in such circumstances, consider a simplified situation in which we wish to fit the two-step profile shown in Fig. 5 with a simple top-hat model with three free parameters: its left and right edges, and its height s .

Table 1. Parameter values for profile models fitted to the original THINGS global spectra.

1	2	3	4	5	6	7	8	9	10	11
Name	v_{ctr} km s ⁻¹	Δv km s ⁻¹	S Jy km s ⁻¹	Δv_{rand} km s ⁻¹	f	α	J %	$\Delta S/S$ %	W_{20} km s ⁻¹	$\Delta W_{20}/W_{20}$ %
DDO 154	375.69±0.12	81.24±0.46	82.10±0.33	9.04±0.15	0.2600±0.0164	-0.0901±0.0081	1.0i	0.0±0.6	103.5	-0.4
DDO 53	18.83±0.20	0.08±0.18	19.88±0.20	12.37±0.14	0.5029±0.2654	-0.0005±0.5241	0.0i	0.4±1.4	44.4	-3.5
Ho II	157.27±0.10	48.97±0.97	218.86±0.73	9.19±0.22	0.2449±0.0486	-0.0663±0.0090	0.5	0.3±0.5	71.0	0.3
Ho I	143.11±0.86	25.70±3.68	40.00±0.33	8.83±0.42	0.5369±0.2501	-0.6686±0.1627	1.6i	0.2±1.1	41.3	-0.3
IC 2574	50.06±0.07	97.84±0.38	385.97±0.64	13.08±0.10	0.3410±0.0099	-0.0788±0.0038	3.4	0.2±0.2	128.6	-0.9
M81 dwA	112.84±0.22	0.08±0.18	4.21±0.08	8.83±0.19	0.5001±0.2662	0.0124±0.5316	0.0	0.6±2.4	31.7	0.5
M81 dwB	345.86±0.96	42.27±4.75	3.74±0.10	8.33±1.08	0.3803±0.2460	0.1226±0.1059	6.9i	1.0±3.7	60.4	3.1
NGC 628	659.36±0.12	45.47±0.20	297.88±0.89	11.88±0.10	0.0080±0.0068	-0.1235±0.0091	2.6	1.5±0.4	75.6	1.6
NGC 925	552.93±0.17	191.24±0.56	229.62±1.05	12.06±0.24	0.2731±0.0114	0.0603±0.0074	6.9	1.0±0.6	222.3	0.5
NGC 1569	-71.38±3.77	73.00±9.28	86.04±1.21	27.65±1.25	0.1929±0.1735	-0.8676±0.0963	2.2	2.3±1.9	123.2	0.2
NGC 2366	100.08±0.07	94.84±0.42	232.91±0.46	9.43±0.12	0.5250±0.0107	0.0767±0.0047	2.7	0.1±0.3	115.1	0.2
NGC 2403	133.44±0.05	228.75±0.12	1047.36±1.23	8.91±0.05	0.3580±0.0022	0.1191±0.0021	4.4	0.7±0.2	251.1	-0.9
NGC 2841	632.49±0.19	566.35±0.44	181.87±0.74	14.62±0.17	0.1228±0.0060	0.0065±0.0049	12.9	0.7±0.6	606.4	-0.2
NGC 2903	555.39±0.12	353.80±0.28	229.88±0.63	12.71±0.14	0.1224±0.0043	-0.1072±0.0044	5.6	1.1±0.4	386.9	0.1
NGC 2976	4.96±0.47	133.88±2.33	44.99±0.54	10.80±0.75	0.5288±0.0486	-0.2219±0.0232	3.6	1.0±1.6	156.2	0.4
NGC 3031	-38.67±0.09	366.28±0.23	1155.72±1.45	23.62±0.10	0.0702±0.0024	0.2966±0.0020	11.2	1.6±0.2	423.8	0.8
NGC 3077	-23.58±0.36	117.12±0.78	257.06±0.79	17.13±0.10	0.0009±0.0008	0.9943±0.0039	7.3	0.2±0.4	117.3	0.0
NGC 3184	593.79±0.14	117.17±0.37	105.37±0.51	9.60±0.17	0.1042±0.0119	-0.0798±0.0074	0.8	0.1±0.7	142.2	-0.6
NGC 3198	661.11±0.07	283.65±0.16	224.91±0.44	10.58±0.08	0.1201±0.0031	0.0353±0.0030	4.1	0.8±0.3	312.2	-0.4
NGC 3351	778.93±0.24	255.37±0.50	49.54±0.48	8.05±0.27	0.0051±0.0040	-0.0001±0.0141	3.8i	1.2±1.3	277.6	-0.3
NGC 3521	798.22±0.10	415.67±0.29	297.23±0.66	19.17±0.11	0.0975±0.0039	-0.0461±0.0034	7.9	0.1±0.3	467.1	-0.3
NGC 3621	729.07±0.08	257.31±0.20	679.58±1.26	11.34±0.09	0.2857±0.0034	-0.0272±0.0031	3.1	0.0±0.3	287.3	-0.4
NGC 3627	720.58±0.73	327.58±2.15	40.07±0.47	22.07±0.93	0.1764±0.0249	0.1190±0.0189	5.3i	1.4±1.6	383.8	1.0
NGC 4214	292.74±0.07	56.34±0.12	200.23±0.39	12.40±0.05	0.0043±0.0035	-0.0458±0.0046	2.3	0.0±0.3	88.9	-0.3
NGC 4449	200.43±0.56	133.64±3.68	263.17±1.15	20.29±0.53	0.9261±0.0487	-0.0222±0.0271	8.5	0.3±0.6	153.5	-3.8
NGC 4736	309.36±0.27	195.56±0.50	76.78±0.49	15.47±0.25	0.0016±0.0013	-0.0500±0.0104	10.7	1.8±0.9	237.2	0.5
NGC 4826	408.87±0.54	288.68±1.22	40.44±0.59	11.79±0.56	0.0110±0.0087	0.3199±0.0223	5.5	2.8±2.0	316.1	0.3
NGC 5055	497.74±0.16	357.96±0.58	376.67±1.09	17.09±0.22	0.4021±0.0065	0.1099±0.0050	7.0	0.6±0.4	400.2	0.4
NGC 5194	455.88±0.28	148.64±0.81	165.24±0.67	17.36±0.31	0.1259±0.0145	0.3317±0.0081	5.3	2.1±0.6	187.8	1.1
NGC 5236	506.81±0.17	172.90±0.33	360.42±0.79	31.73±0.14	0.0030±0.0024	0.1650±0.0045	4.4	1.0±0.3	253.6	-5.4
NGC 5457	228.74±0.09	138.68±0.41	1093.04±1.39	23.15±0.11	0.0134±0.0083	0.2260±0.0026	5.0	0.9±0.2	195.8	-0.3
NGC 6946	45.03±0.13	210.84±0.36	508.91±1.71	11.09±0.14	0.3078±0.0068	0.1004±0.0057	9.3	1.2±0.5	238.9	0.1
NGC 7331	816.52±0.26	479.34±0.57	178.52±0.98	15.07±0.27	0.0414±0.0087	0.1150±0.0080	2.1i	0.2±0.8	518.8	0.2
NGC 7793	226.75±0.11	157.62±0.40	245.89±0.80	12.43±0.15	0.1312±0.0092	-0.0601±0.0054	5.7	0.4±0.5	190.3	0.1

Notes. Columns are as follows. 1: source name. 2-7: the 6 model parameters, being mean values of the MCMC distribution. 8: wiggle fraction J as defined in equation 8. 9: fractional difference between data and fitted values of total flux, also plotted in Fig. 4. 10: width at 20% of peak flux density of the fitted profile. 11: fractional difference between data and fit values of W_{20} , also plotted in Fig. 6.

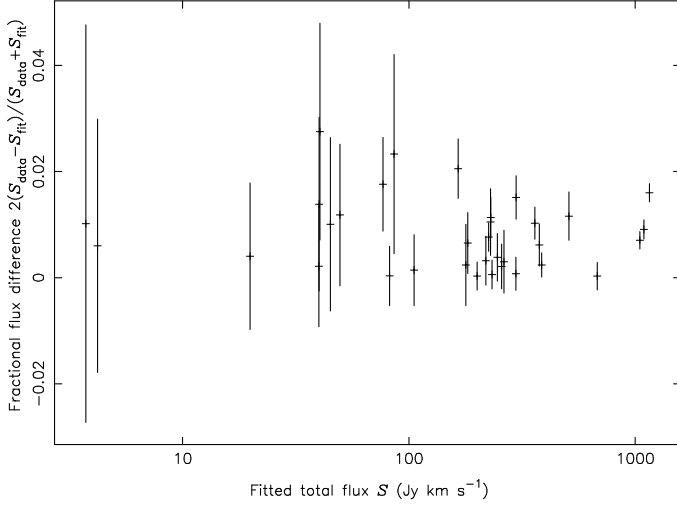


Fig. 4. Fractional difference between the flux S_{data} , derived from summing the flux densities in each data channel, and the value S_{fit} , the equivalent sum of flux densities from the fitted line-profile model. This fraction was calculated from the expression $2(S_{\text{data}} - S_{\text{fit}})/(S_{\text{data}} + S_{\text{fit}})$. The x coordinate is the fitted flux S given in Col. 4 of Table 1, whereas the fraction itself is given in Col. 9. S_{fit} is identical to S except in the cases of DDO 53, NGC 2976, NGC 3077 and NGC 6946. For these galaxies, some channels around velocity zero were omitted from the calculation of both S_{data} and S_{fit} . Uncertainties were calculated from the two independent contributions using the standard propagation formula.

Obviously there is nothing to be gained by moving the right edge of the model away from v_2 ; and given any two values for the edge locations, determining the best-fit height is trivial. The only open question is where within the range $[v_0, v_1]$ should we place the left edge v of the profile such that χ^2 is minimized. We define a kind of non-discrete, noiseless analog of χ^2 as

$$\chi^2 = (v_1 - v)(s - s_0)^2 + (v_2 - v_1)(s_1 - s)^2.$$

After optimizing the model height s this becomes

$$\chi^2 = (s_1 - s_0)^2 (v_2 - v_1) \frac{v_1 - v}{v_2 - v}.$$

Clearly this has its lowest allowed value when $v = v_1$. But this means that the profile is not fitting the pedestal at all, and the data then has more flux than the model, which is consistent with the characteristic upward deviations seen in Fig. 4. Pedestals are noticeable in 5 of the galaxy profiles shown in Figs. C.1 to C.34, namely for Ho II, NGC 628, NGC 5194, NGC 5236 and NGC 5457. Of these, all but Ho II are among the group of 6 worst cases of flux underestimation.

Figure 6 compares the velocity widths at the 20% height of the fitted profiles (W_{20}) to the same figure calculated from the raw data values. For both model and data, the 20% flux density level was calculated with respect to the maximum value for that spectrum; the velocity at this level on either wing of the profile was calculated by linear interpolation between the most distal pair of velocities which straddled the 20% level. This simple technique is only accurate for the data because of the relatively high S/N of the THINGS observations. Some discussion of difficulties which arise in linewidth estimation when the spectrum is noisy is given in Sect. 3.2.2.

The agreement between the fitted width and the widths from the data is seen to be very good. In few cases is the difference larger than the 1% level. The anomalously high value ($W_{20,\text{fit}} > W_{20,\text{data}}$) belongs to the M81 dwarf B; the three low

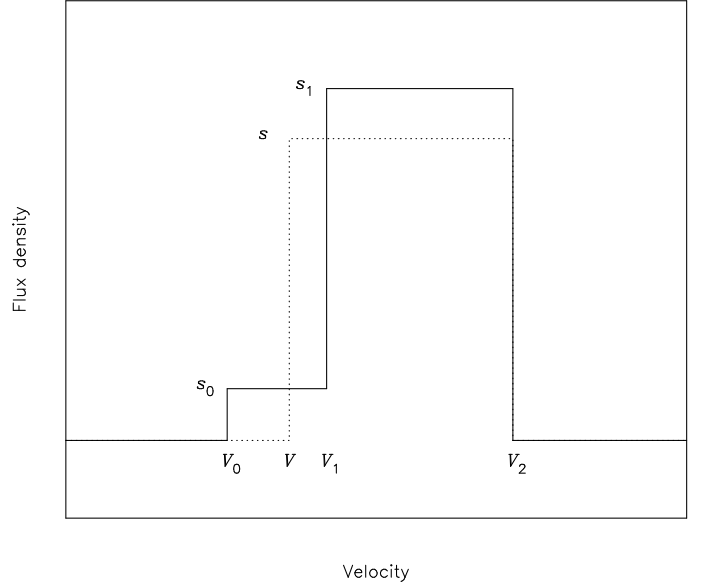


Fig. 5. Demonstrates a mechanism by which the flux under a spectral line may be underestimated. The solid line shows a simple model of a continuous spectrum which has a pedestal extending some way to the left. The model, indicated by the dotted line, has itself no matching pedestal. χ^2 is minimized if $v \rightarrow v_1$ and $s \rightarrow s_1$.

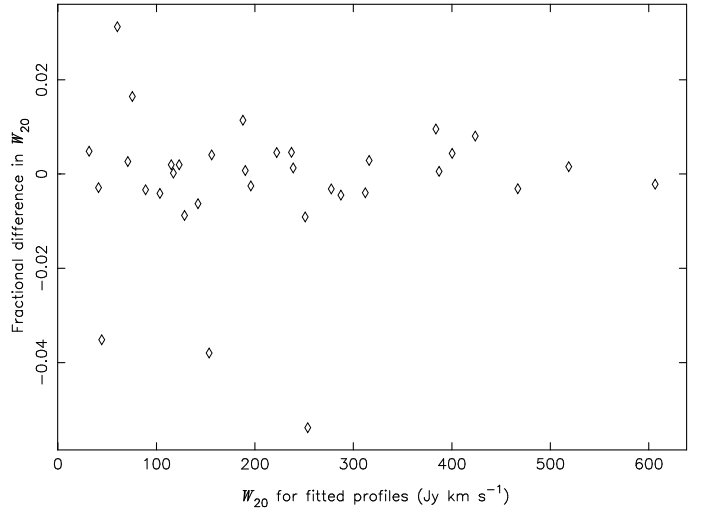


Fig. 6. This plot compares the linewidth $W_{20,\text{data}}$ at the 20% height estimated from the observed THINGS spectral line to $W_{20,\text{fit}}$, the same value for the profile fitted in the present paper to that line. The fractional linewidth difference was calculated from the expression $2(W_{20,\text{fit}} - W_{20,\text{data}})/(W_{20,\text{fit}} + W_{20,\text{data}})$. Values for $W_{20,\text{fit}}$ are given in Col. 10 of Table 1, and the fractional difference itself is given in Col. 11.

values ($W_{20,\text{data}} > W_{20,\text{fit}}$) to DDO 53, NGC 4449 and NGC 5236. For M81 dwB and DDO 53, the width of the spectrum channels is several percent of the W_{20} width, and therefore entirely accounts for the error. In NGC 4449 and 5236 the profile model is clearly seen not to be a good fit. However, we note that NGC 4449 is an interacting galaxy with a significant amount of H I outside the disk. And for NGC 5236 the THINGS observations have many missing spacings, which may explain the curious steps in its line wings, which are probably the cause of the inflated value of $W_{20,\text{data}}$.

De Blok et al. (2008) fitted a tilted-ring model to a 19-member subset of the THINGS galaxies and obtained high-

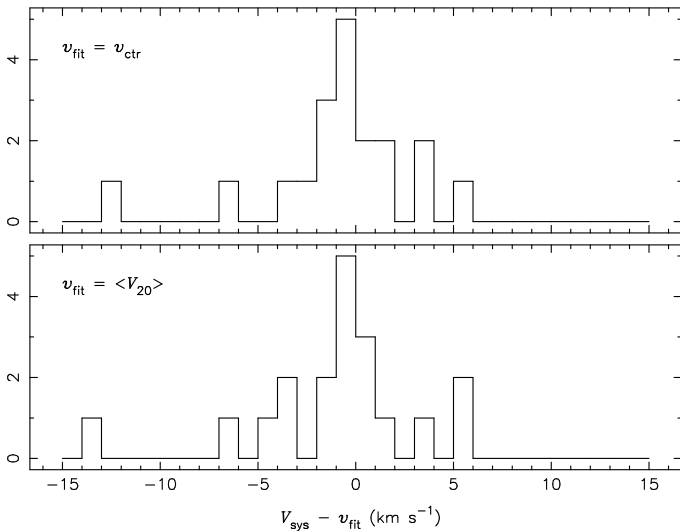


Fig. 7. Frequency histograms of the difference between values of systemic velocity from two different sources. The first value, V_{sys} , is that given in Col. 6 of Table 2 of de Blok et al. (2008). These values were derived from tilted-ring fits to rotation curves of a 19-member subset of THINGS galaxies. The second value, v_{fit} , was calculated in the present work. In the upper plot, the line centre parameter v_{ctr} of the fitted profile was used for v_{fit} ; in the lower plot, v_{fit} is the mean of the low and high velocity values at the 20% height level on the fitted profile.

precision H I rotation curves for these. In Fig. 7 we compare their systemic velocities with two measures of line centre from the fitted profile, namely the fitted line centre parameter v_{ctr} , and the velocity $\langle V_{20} \rangle$ obtained from the mean of the 20% velocities. The systemic velocities V_{sys} are taken from Col. 6 of Table 2 of de Blok et al.. With exception of the outlier at about -14 km s^{-1} , which is NGC 3627, the spread in differences is centred on zero with a standard deviation on the same order as the channel width for these galaxies, which was about 5.2 km s^{-1} for 12 out of the 19, half that for the rest.

De Blok et al. found that a V_{sys} calculated from the global line wings of NGC 3627 returned values in the range 717 to 720 km s^{-1} , which fall much closer to our value of 720.6 . NGC 3627 appears to have kinematic asymmetries, such as have been described and discussed by Swaters et al. (1999). However, in the two galaxies studied by Swaters et al., the disturbances seem to be confined to the inner velocity regions. Disturbances in the inner-disk kinematics don't affect the rise and fall of the global profile, and thus should have only a small effect on the V_{sys} derived via fitting the present model. In contrast to this, for NGC 3627, Fig. 79 of de Blok et al. shows a significant excess of gas at velocities 10 km s^{-1} or more greater than their tilted-ring fit, right at the trailing edge of the velocity distribution. The present model is sensitive to such distortions, so this is arguably the reason for the inconsistency between our value of V_{sys} for this galaxy and that of de Blok et al. (2008).

Andersen & Bershadsky (2009) found that line centres obtained via a moment analysis were biased in proportion to the asymmetry as measured by the 3rd moment. Since asymmetry is built in to our model from the start, we would not expect it to suffer from a similar problem. Figure 8 confirms that this is the case. This figure is analogous to Fig. 8 of Andersen & Bershadsky. Despite that both our velocity scatter and uncertainties are smaller, no correlation is apparent in our Fig. 8.

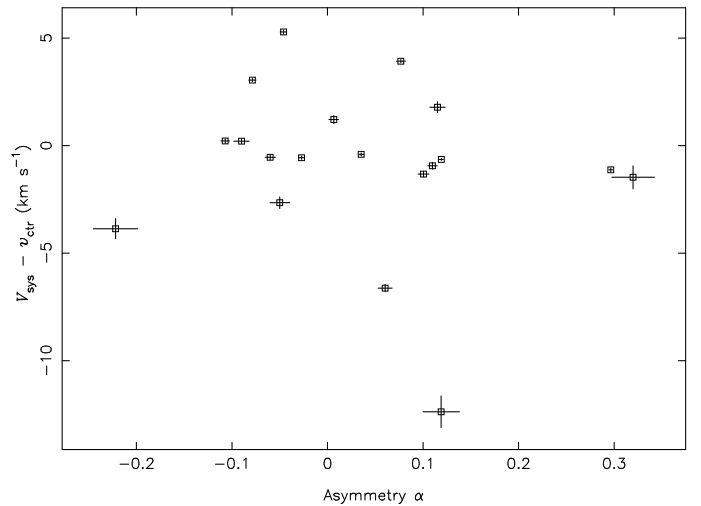


Fig. 8. Scatter plot comparing the asymmetry parameter α of the fitted profile model to differences between the present fitted centre velocity v_{ctr} and the systemic velocities V_{sys} obtained by de Blok et al. (2008). The velocity difference is the same as that histogrammed in the upper panel of Fig. 7. Only the 19-member subset of THINGS galaxies examined by de Blok et al. are shown.

3.2. Fitting to coarsened and noisified data

3.2.1. Introduction

We argue in the introduction that the main use for this profile model is to extract information in a systematic way from the kinds of H I spectra most commonly encountered in blind surveys, namely with low S/N, and a channel width of about 10 km s^{-1} . In order to test how well the model serves this purpose, data is needed with *a priori* known values of line width, total flux etc. to allow comparison with the values obtained from fitting the model. Also a large number of test spectra is desirable to reduce or at least ascertain the uncertainties in the results. The raw THINGS spectra will not serve for this: they are too few in number, have unknown *a priori* values, have too high S/N, and significantly higher spectral resolution than we normally expect to encounter in blind H I surveys.

An alternative would be simply to simulate the required numbers of noisy, coarsely-binned profiles. However such a simulation always carries with it some uncertainty about how applicable its results are to real measurements. Simulations are also vulnerable to the criticism that one only gets out only what one puts in.

We can however make use of the THINGS profiles to generate realistic data simply by binning them into wider velocity channels and adding a lot more noise. Lewis (1983) made use of a similar method in investigating line width measurement biases. This is also the approach which is taken in the present section.

What we do here is run several Monte Carlos, in each of which an ensemble of test spectra is generated. The line-profile model is fitted to each spectrum in an ensemble, and the width, centre and total flux of the line are also estimated using traditional direct methods. For each ensemble, and for each property of interest (line width etc.), an average is formed from the values obtained from the fitted profiles on the one hand and the direct measurement on the other. The input values of the properties are known and so are available for comparison. The aim is to demonstrate that the average value of each property is better estimated via model fitting than by direct measurements from the spectrum.

3.2.2. The Simulation Monte Carlo

Five of the THINGS galaxies were selected, the sample being chosen so as to cover a range of shapes and sizes. These five are the ones shown in Fig. 2. They all had channel widths of about 2.6 km s^{-1} , except for NGC 4214, for which the width is half this. For each of the five, a Monte Carlo ensemble of 100 spectra was generated for each of a set of twelve S/N values, the S/N figure being calculated according to the ALFALFA formula, which is discussed in Sect. 3.2.3. The twelve values were evenly spaced in a logarithmic sense over a range between about 2 and 100. Different amounts of S/N offset were applied to the five galaxies to prevent overlay of graph points when results from several galaxies appear on the same figure.

As already mentioned, the desired channel width in this Monte Carlo was about 10 km s^{-1} . This was most simply achieved by averaging an integer number of the channels of the original spectra, that number being 4 for all but NGC 4214, for which a factor of 8 was used. This scheme also permitted dithering of individual spectra by random offsets in the range from zero to one less than the rebinning factor. This dithering was to avoid systematic effects from the centre velocity of the input spectrum always having the same relation to the channel boundaries in the rebinned scheme.

3.2.3. Calculating the S/N

Signal-to-noise ratio can be defined most simply as the ratio between the maximum or peak height of the input spectrum and the noise standard deviation or RMS. However this does not correlate well with the detectability of H I lines, because a broader line is usually easier to detect than a narrow one having the same peak-to-RMS S/N. For this reason it is usual to define some alternative measure which takes account of the line width. The ALFALFA survey makes use of the following formula (equation 2 of Haynes et al. 2011):

$$S/N_{\text{ALFALFA}} = \frac{S}{\sigma W_{50}} \sqrt{\frac{\min(400, W_{50})}{2 \Delta v_{\text{chan}}}}$$

We used the same to calculate the S/N values for our Monte Carlo.

Note firstly that, for any given profile shape, $S/N_{\text{peak}} \propto S/N_{\text{ALFALFA}}$, but the proportionality constant differs from profile to profile. Note secondly that ALFALFA take $S/N_{\text{ALFALFA}} > 6$ as their detection criterion (Haynes et al. 2011).

A sample of a spectrum produced in the Monte Carlo is shown in Fig. 9. This has a S/N equal to 3.9, which is a little below the ALFALFA detection criterion. As such it represents a very typical example of the quality of profile commonly seen in such surveys.

3.2.4. Fitting to the data and analysing the results

The simplex algorithm was used to fit each spectrum. Several fitting runs were done with increasing values of maximum permitted number of iterations, and decreasing values of the convergence criterion, in order to be sure that the fits were properly converging. The posterior described in Sect. 2.4.1 was the object function fitted, but this time some attention was paid to priors. Details of these and their justification can be found in appendix B.2.

There are two ways one can analyse such data: one can compare the fitted values of the model parameters to those values

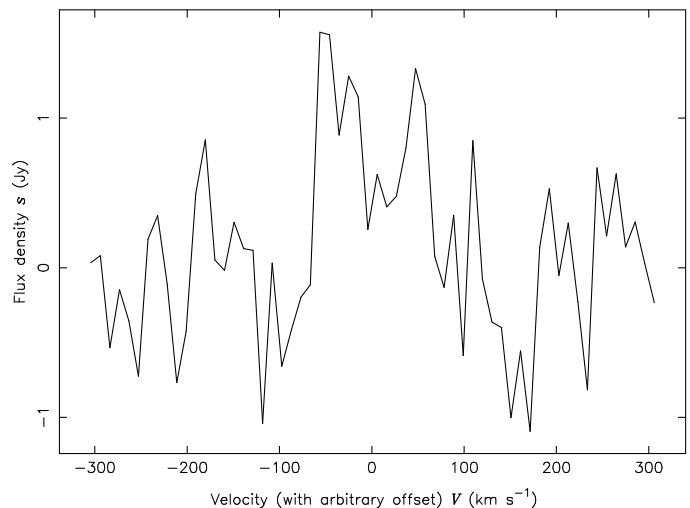


Fig. 9. An example of a rebinned and noisified profile such as described in Sect. 3.2. The template spectrum was the THINGS observation of NGC 3184. The template data have been rebinned by a factor of 4 to give a channel width of 10.35 km/s ; noise has then been added to bring the signal-to-noise ratio to a figure of 2.0 in peak-to-RMS terms, which for this galaxy is equivalent to about 3.9 according to the ALFALFA formula, as described in Sect. 3.2.3.

obtained from fits to the original THINGS template spectra; and one can compare measurements derived from fitting a model to those obtained via more traditional, non-parametric methods. Both sorts of analysis have been done for the ‘coarse’ data. The corresponding results are described respectively in Sects. 3.2.5 and 3.2.6.

It is always going to be difficult to obtain reliable estimates of line parameters from data at such low S/N values as in Fig. 9. There are three quantities which are of central importance: the total flux under the spectral line, its central or systemic velocity, and its width.

In the case of total flux, a typical non-parametric estimation method is that described in Haynes & Giovanelli (1984): i.e. simply to ‘integrate over the observed signal’. Implicit in this however is a knowledge of where the ‘observed signal’ begins and ends. For high S/N signals this is often done by visual inspection. In this case exactitude is not too important - one can afford to be generous, since the added noise per extra channel is small. Judging the line boundaries is of course much more problematic in the weak-signal limit. There are also practical difficulties involved in human inspection of large numbers of spectra.

Line width and line centre are often derived from the same source, namely two measurements of the low- and high-velocity edges of a line. Line width is taken from their difference and line centre from their average. Bica & Giovanelli (1986) present several typical methods of estimating the edge velocities. These all start by deciding a flux density value, then interpolate linearly between adjacent channels which straddle this value. The flux density chosen is some percentage of a characteristic value for the line as a whole, which may be the mean flux density, or the maximum within the line profile, or some function of the peaks at the horns of the profile, in cases where these can be measured. In all cases such values for the characteristic flux density depend upon either the channel width or on implicit assumptions about where the line begins and ends.

The largest blind H I survey (still on-going) is ALFALFA (Giovanelli et al. 2005); their most recent catalog release paper is Haynes et al. (2011). The latter authors describe their edge-

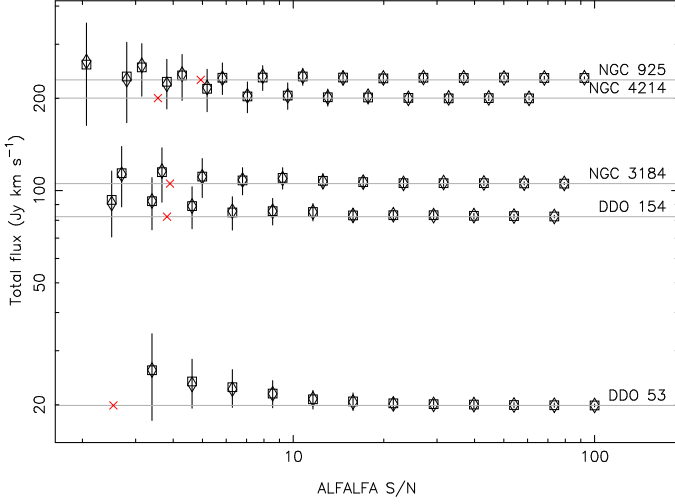


Fig. 10. Mean total flux of an ensemble of fitted line profiles over a range of values of ALFALFA signal-to-noise ratio (S/N). Only 5 of the THINGS galaxies are shown, as labelled. Construction of these Monte Carlos is described in Sect. 3.2.2. The square symbols give the mean, the diamonds the median, and the vertical bars show the standard deviation for each ensemble. The half-tone lines give the values from the profile fits to the original THINGS data. Red Xs show the points at which the peak-to-RMS S/N equals 2.

estimation procedure as similar to that of Springob et al. (2005), in which polynomials (in practice nearly always ‘polynomials of order 1’, i.e. straight lines) are fitted to several channels at the line rolloff on each side, the interpolation flux density being 50% of the horn height on that side. In practice this is not much different to schemes already mentioned because, unlike the profiles shown in Fig. 2 of Springob et al. (2005), in ALFALFA the effective channel width of 10 km s⁻¹ would not provide many channels in the rolloff.

To settle on a point of comparison therefore we decided to employ, as our ‘traditional, non-parametric method’ of estimating the line edges, the simple scheme as follows:

1. Determine s_{\max} , the highest flux density within the spectrum.
2. Identify the number j_{lo} of the lowest-velocity channel whose s value exceeds $0.5s_{\max}$.
3. Identify the number j_{hi} of the highest-velocity channel whose s value exceeds $0.5s_{\max}$.
4. Interpolate between $j_{\text{lo}} - 1$ and j_{lo} to obtain v_{lo} , and between j_{hi} and $j_{\text{hi}} + 1$ to obtain v_{hi} .

It is thus a maximizing algorithm.

3.2.5. Monte Carlo results I: comparison with high-S/N fits

Some of the more important results of the Monte Carlo described in Sects. 3.2.2 to 3.2.4 are shown in Figs. 10 to 15.

In Fig. 10, the variation of the mean value of total flux fitted to the spectra is shown against the S/N, for all five galaxies in the trial. Some upward bias in the values becomes visible at low S/N values, but this always remains within 1 standard deviation of nominal.

Figure 11 is a similar plot, this time for the intrinsic line width. DDO 53 is not shown, because line width is small and poorly constrained for this galaxy, even in fitting to the raw THINGS spectrum. Similar slight biases are observed in three of the four galaxies; it is not clear why NGC 4214 seems to be so badly affected at low S/N.

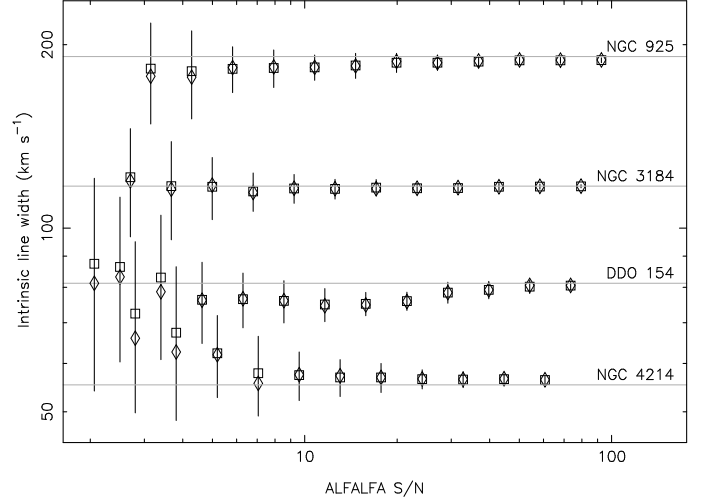


Fig. 11. Similar to Fig. 10 except that the fitted value of intrinsic line width is shown instead of total flux. DDO 53 is not shown because the intrinsic line width for this galaxy is not significantly different from zero.

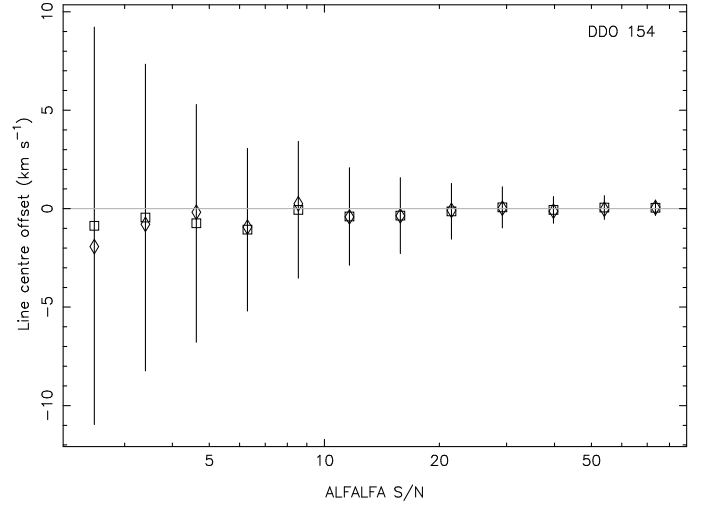


Fig. 12. Mean fitted line centre offset is shown for DDO 154, at the same values of S/N as in Figs. 10 and 11. No other galaxies in this subset are shown because all 5 show similar characteristics.

It is important to note that, at the highest S/N levels shown, the linewidth clearly asymptotes to the value measured in Sect. 3.1 and tabulated in Col. 3 of Table 1, despite the fact that the spectral resolution of the present data is several times poorer than the original THINGS observations. This emphasises that no ‘instrumental broadening’ correction to linewidth measures is necessary when fitting a line-profile model.

The curious under-estimation by up to 2σ of the linewidth for DDO 154 at quite large values of S/N remains unexplained, although it is worth noting that it is never more than about 8 km s⁻¹, which is less than the channel width of these spectra. In the next section it is demonstrated however that for DDO 154 the W_{50} linewidth remains correct in this range - so clearly the decrease in the fitted ‘intrinsic’ value is compensated by an increase in the ‘turbulent broadening’ value.

Note that the mean fitted linewidth for all five galaxies remains accurate down to the ALFALFA limit of detectability.

The mean and standard deviation of the line centre offset (that is, the difference between the fitted line centre and the in-

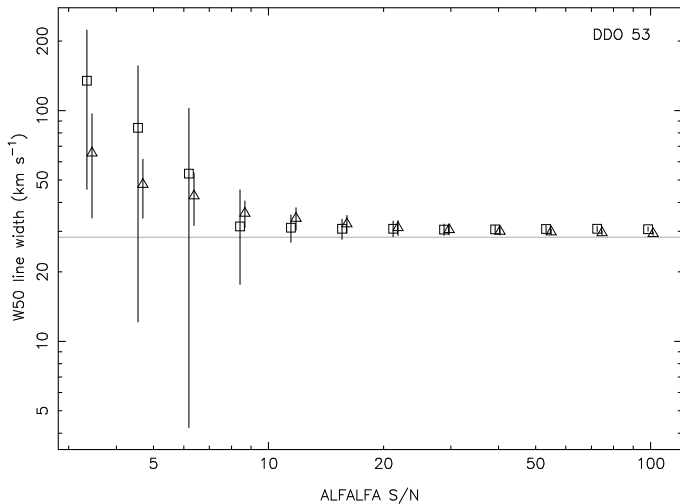


Fig. 13. For the same range of S/N values as in Figs. 10 to 12, line width at the 50% height for the raw spectrum versus the same for the fitted profile is compared. DDO 53 is shown here. The square symbol indicates the mean value for the raw spectra, the triangle the mean for the fitted profiles. These respective points have been slightly offset horizontally for clarity. The half-tone line indicates the value for the profile fitted to the original THINGS data.

put value) are displayed in Fig. 12 just for DDO 154. All the galaxies show similar results. No significant bias is seen, and the scatter remains within the channel width for ALFALFA S/N values greater than 5.

From these results we can conclude that, so far as the important parameters of the fitted profile model go, there is little bias evident when the spectral resolution and S/N are lowered to typical levels expected in blind surveys.

3.2.6. Monte Carlo results II: comparison with non-parametric methods

Figures 13 through 15 all make the same comparison, but a single galaxy is plotted per figure for clarity. The other two galaxies have not been shown because their results are similar. For each value of S/N in each plot, two values of W50, the line width at 50% of maximum height, are shown: the error bar centred on a triangle comes from measuring the fitted profile; that centred on the square, from the raw spectrum, using the simple algorithm discussed in Sect. 3.2.4. It is obvious from all plots that finding W50 from the fitted profile yields a more accurate result. The effect is most stark for broad profiles, which for the same ALFALFA-formula S/N have a much lower peak-to-RMS value.

3.2.7. Discussion

Lewis (1983) made a similar study of the effect of noise on measurements of H I line position and width. Although Lewis's study was in many respects more comprehensive than the present one, this author applied traditional direct methods to measure the positions of line edges, and deprecated the use of functions fitted to the line shoulders over several channels. Lewis argued that the number of channels spanning the line shoulder was of the same order as the minimum number of parameters required by any useful fitting function, and therefore that fitting was equivalent merely to interpolation, and could therefore not be expected to yield any improvement in precision.

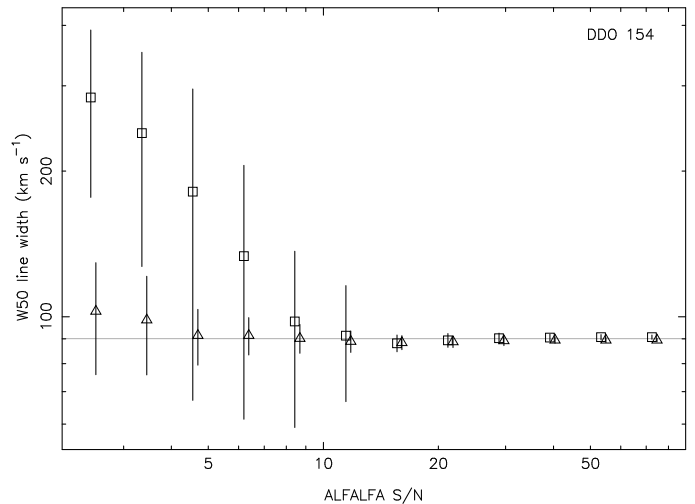


Fig. 14. Same as for Fig. 13 except DDO 154 is shown.

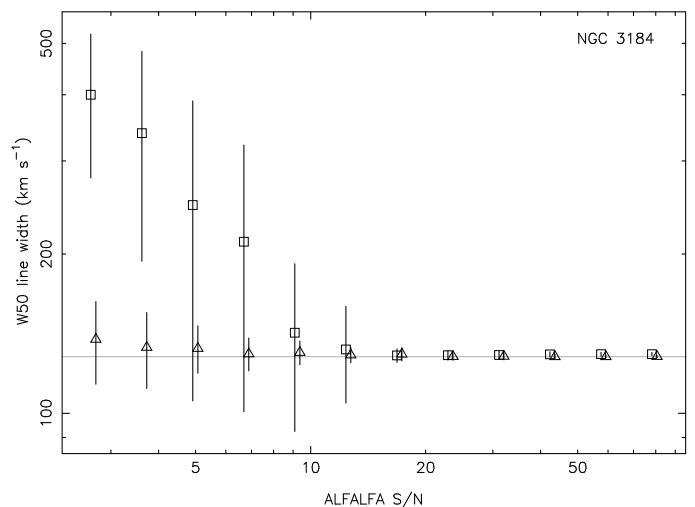


Fig. 15. Same as for Fig. 13 except NGC 3184 is shown.

In opposition to this view is the Bayesian probability theory, which says that the most complete knowledge possible of the probability distribution of linewidth values is obtained by use of Bayes' theorem in conjunction with a model of the spectral line. A seeming oversupply of model parameters is no objection provided the unwanted ones are marginalized out. The only necessary criterion is that the model reproduces accurately the widths of real spectral lines. For the present model, this is well demonstrated in Sect. 3.1.3.

Nevertheless it seems reasonable that a model with 6 parameters will be poorly constrained if the spectral line extends over no more than about the same number of velocity channels. It is of interest therefore to consider the likely shape of the Bayesian posterior distribution in the extreme case that the spectral line has most of its flux in a single channel. In this case it seems clear that the line centre and total flux parameters will remain well-constrained. On the other hand, the fraction-solid and asymmetry parameters can be expected to be completely unconstrained, with the posterior density remaining about the same if they are varied through their respective ranges (as for example in Fig. B.5). Since the Gaussian spread of the spectral line is modelled by the Δv_{rand} parameter, its value will be most tightly constrained by the flux densities in the channels adjacent to the central one. If these are insignificant, then this parameter will be unconstrained

within a range between zero and about the width of a channel. The final parameter, the intrinsic linewidth, will be expected to have the same character (this again is similar to what is seen in Fig. B.5).

In fact none of these indeterminacies and degeneracies matter, because in practice one integrates the posterior over all the dimensions (i.e. parameters) of no present interest. For example if one wants a probability distribution of W_{50} for a given spectrum, this is very easy to obtain from the Bayesian formulation for the respective data. The number of parameters in the model should be viewed merely as a detail of the way the problem is worked out, having no importance for the final result.

Another issue which deserves some discussion concerns the biases in fit parameters observed in Figs. 10 to 15. If the Bayesian formulation is optimum, one might ask, why does it still lead to biased results? The answer is that the Monte Carlo procedure adopted in the present section is not the way one ought to proceed with real data. In no case is it optimum to fit separate spectra, then average the resulting parameter values. This is done here only because we have no option if we wish to make comparison with standard methods of estimating line parameters directly from the data. If one did not have the necessity of comparison with non-Bayesian methods, the correct approach depends on the nature of the quantity to be estimated. Either we have an ensemble of observations of the same object, or an ensemble of observations of different objects. In the former case, the correct Bayesian approach is to fit once to the entire data set. In this case we are back (given sufficient data) in the high S/N regime and any biases will be correspondingly insignificant. If the objects are all different, then the correct approach is hierarchical modelling. An interesting recent example of this treatment is Brewer et al. (2014). Essentially one defines a hyper-model which describes the distribution of values of properties across the ensemble of objects, and tries to constrain its hyper-parameters. Although any individual spectrum is low S/N, again a sufficiently large data set will eventually reach the high-S/N, thus low-bias regime as regards the hyperparameters.

3.3. The THINGS galaxies in EBHIS

3.3.1. Overview

The semi-simulated spectra generated in Sect. 3.2 still fall some way short of real life and therefore cannot be expected to present the full range of difficulties one encounters in trying to analyse real observations. We wanted also to test the profile model on single-dish data. The EBHIS survey (Effelsberg Bonn H I Survey, Winkel et al. 2010; Kerp et al. 2011) is convenient for this, since its survey coverage has included most of the THINGS galaxies, yet as a single-dish survey, it necessarily has much reduced spatial resolution when compared to the interferometer used by Walter et al. (2008) to make their THINGS observations. An extensive comparison of the EBHIS observations of the THINGS galaxies might be interesting and useful, but is beyond the scope of the present paper. Here we just wish to show that profile-fitting works well also under less than ideal conditions. Specifically we are interested in the following questions:

- Can we cope with non-flat baselines and interference?
- Can profile-fitting at separate pixels across the source return a map of the source flux which is better (less noisy) than one obtained simply by adding flux densities over a range of channels?

- Can the variation in the asymmetry parameter across the source give clues to the galaxy orientation and inclination, even with poorly-resolved sources?

3.3.2. Baseline fitting

The baseline problem differs between single-dish vs. interferometer spectra. Single-dish measurements tend to be worse affected by Fabry-Perot type effects which can produce waves in spectral baselines. Spectra made with an interferometer on the other hand tend to have flat baselines, although a spatially superimposed continuum source can raise and otherwise perturb the baseline. Future, deep interferometric surveys for H I may, however, begin to be affected by source confusion, which may necessitate wider use of baseline fitting.

Traditionally, non-flat baselines have been dealt with by fitting a function, often a simple polynomial, to stretches of channels on either side of the line of interest. The fit is interpolated across the line profile but channels within the profile make no contribution to the fit. This is less than ideal for three reasons: firstly because of the restricted number of channels which contribute information to the fit; secondly because it relies on a decision as to where the profile begins and ends; thirdly, selection of the degree of the fitting polynomial is usually done ‘by eye’.

The approach taken in the present paper is simply to add a baseline function to the existing profile model, and fit the combined model to all selected channels. This also lends itself to a Bayesian approach to determining the best order of the baseline function. This technique is more fully described in appendix B.6.

Chebyshev polynomials have been preferred for the baseline function in the present paper, because they are easy to generate, and more orthogonal and numerically tractable than simple polynomials. Note however that a Chebyshev polynomial is still just a polynomial, and an n th-order Chebyshev fitted to some data will result in exactly the same function as fitting a simple n th-order polynomial would (B. Winkel, private communication). The Chebyshev is just better behaved from a computational point of view.

The EBHIS cubes have 3000 spectral channels spanning nearly 3900 km/s, which is far wider than any of the lines fitted. It isn’t necessary to fit to the full velocity range, so in all cases a small section only, about 3 or 4 times wider than the spectral line, was selected for fitting purposes.

3.3.3. Results

Results are presented here for only a single galaxy, DDO 154.

Figure 16 shows the spectrum of DDO 154 obtained by integrating a 9×9 pixel area of the relevant data cube. The profile shown is the mean of 10^4 iterations of a converged MCMC, but it is almost indistinguishable to the eye from the best-fit profile from a Levenberg-Marquardt optimisation. The baseline is fitted by a sum of Chebyshev polynomials truncated at order 5 (i.e. giving 6 orders in total), which number gave the maximum Bayesian evidence (see appendix B.6).

A spectrum summed across the entire breadth of this cube showed no evidence for narrow-band RFI in any channel, so no channels were excised before fitting.

Because there is a significant baseline contribution in this spectrum, careful attention was paid to baseline priors, which were estimated from the non-source portions of the cube as described in Sect. B.7. As can be seen from the figure, the prior accounts for quite a lot of the baseline over the small spatial ex-

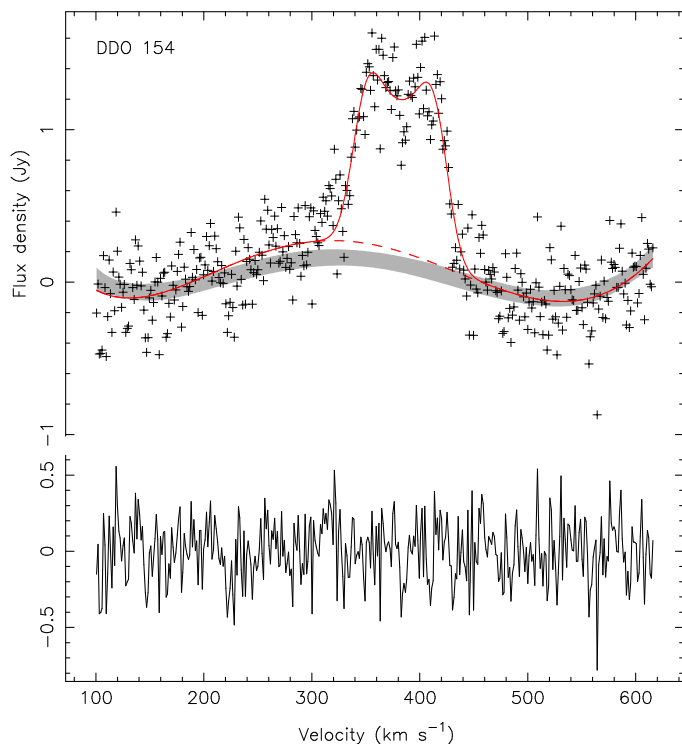


Fig. 16. Constructed from the EBHIS cube for DDO 154. The upper plot shows a spectrum from a section of the cube, integrated over the same spatial dimensions as shown in Figs. 17 to 19. The lower plot shows the residuals after subtraction of the fitted model. The red line shows the MCMC-mean profile model, with 6 orders of Chebyshev function. The solid line shows the whole profile, whereas the dashed line just shows the baseline. The ± 1 sigma range of the baseline prior is indicated by the grey band.

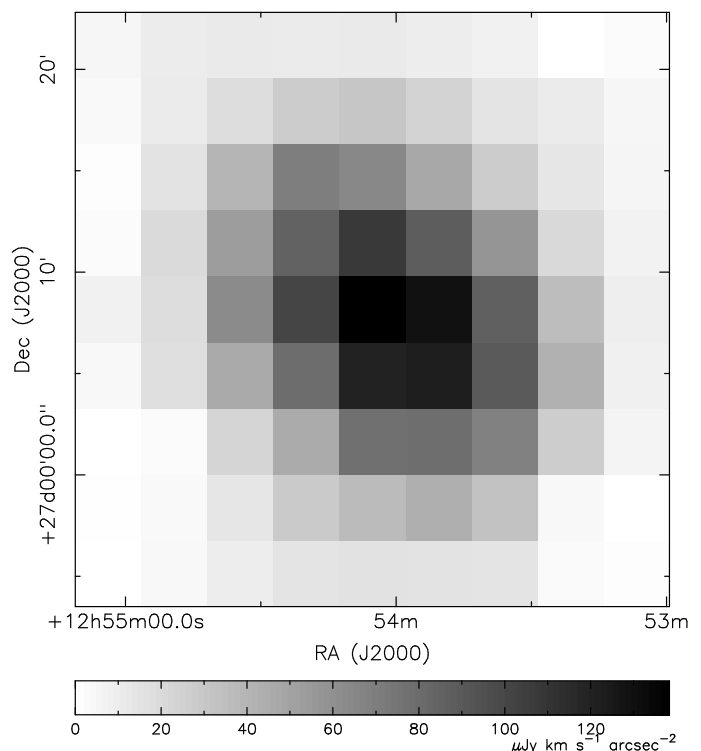


Fig. 17. Constructed from the EBHIS cube for DDO 154. The plot shows the fitted total flux at each of a small range of spatial pixels.

tent of the source. This gives one confidence that the baseline subtraction is accurate, and thus also the fitted total flux of the spectral line.

A Levenberg-Marquardt optimization returned a best-fit value of $102 \pm 4 \text{ Jy km s}^{-1}$ for the total flux of DDO 154, whereas the MCMC returned a mean value of $100 \pm 5 \text{ Jy km s}^{-1}$. These values are significantly higher than the value of 82 fitted to the data of Walter et al. (2008). However, it is known that interferometric measurements can miss flux. Our value agrees very well with the value of $105 \pm 5 \text{ Jy km s}^{-1}$ reported by Carignan & Purton (1998) from a careful combination of interferometer and single-dish data.

Next, the baseline-plus-line model was fitted to spectra extracted at single spatial pixels across the same 9×9 area integrated to produce the spectrum in Fig. 16. 9×9 maps of the fitted parameter values were made. Figures 17 and 18 show two of these, respectively for the total-flux parameter and the asymmetry parameter. (The pixels of Fig. 18 were set to null where the total flux fell below 15% of maximum.) It is of interest to compare these to the much higher-resolution versions in Fig. 53 in Walter et al.. As one might expect, the single-dish EBHIS survey seems to detect more extended structure than Walter et al.'s interferometric measurements.

Clearly visible in the asymmetry map is the orientation of the galaxy's major and minor axes, even though DDO 154 is barely resolved by the 9 arcmin FWHM beam of the Effelsberg telescope at L band (Kerp et al. 2011).

The remaining parameter maps are of less interest and are not shown. The map of the fitted line centre also (as one might

expect) shows indications of the major and minor axes, but the signal is not so unequivocal as with the asymmetry map.

As far as the baseline fits go, no obvious trends are visible either in the 6 maps of the individual Chebyshev coefficients, or

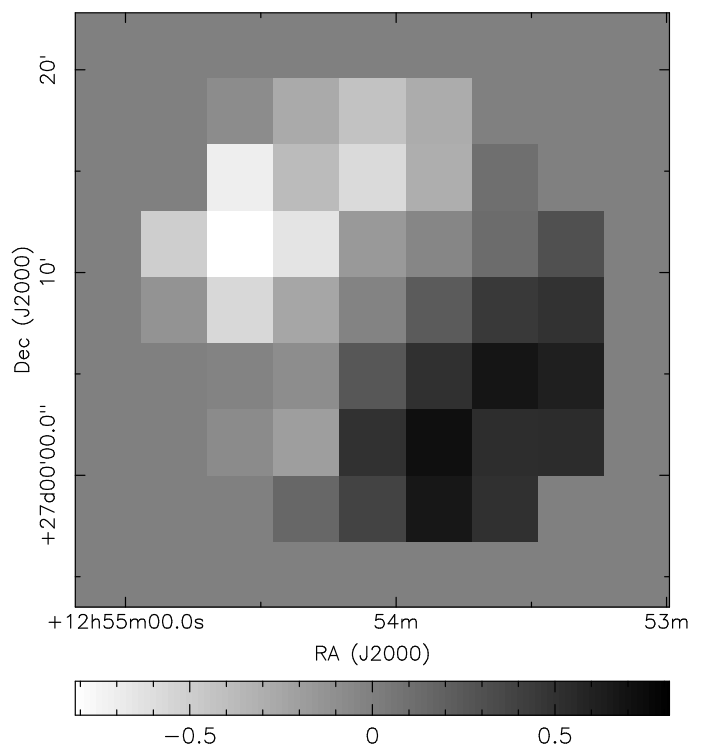


Fig. 18. Same as for Fig. 17 except the asymmetry parameter is shown. Pixels are set to median grey where the total flux drops below 15% of maximum.

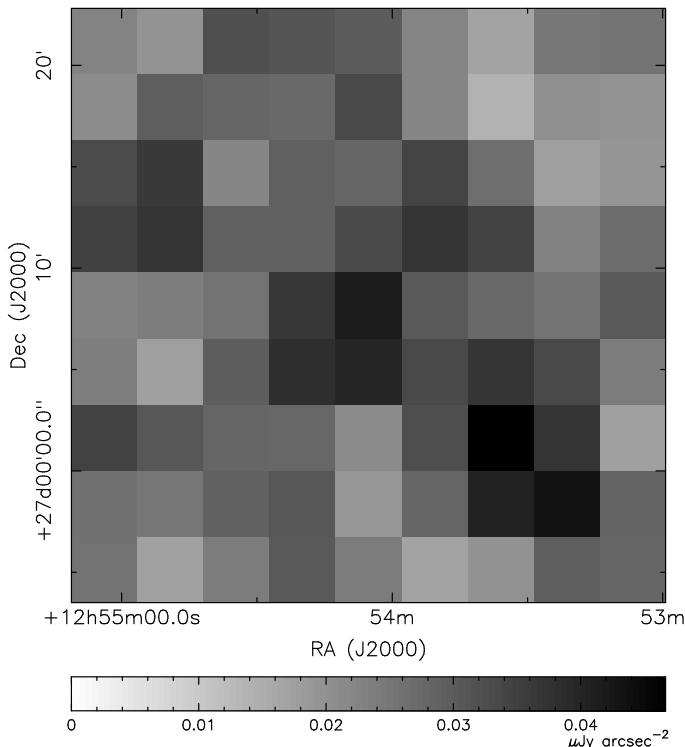


Fig. 19. Same as for Fig. 17 except the RMS of the Chebyshev coefficients is shown.

a map of their RMS value (Fig. 19). There is no bright continuum radio source to disturb the baseline at the location of DDO 154: NVSS (Condon et al. 1998) shows that there is no source with an average flux density at L band greater than 90 mJy/beam within the approximately 30" by 30" square shown in Figs. 17 to 19. The first Chebyshev coefficient (which is just a DC offset) does appear slightly elevated at the centre of the source, but only at the 2-sigma level.

4. Conclusion

The two most desirable measurements of the global H I spectral profile of a galaxy are its width and area. Combined with optical measurements of the inclination and brightness of the galaxy, these allow the galaxy's H I mass and intrinsic optical luminosity, and thence its distance, to be estimated via the baryonic Tully-Fisher relation. The systemic velocity of the galaxy comes next in importance. This quantity, when compared to the Hubble velocity appropriate to that distance, additionally allows one to calculate the peculiar or local velocity of the galaxy in the line of sight to the observer. A profile model is described here which permits one to make accurate estimates of these quantities.

In Sect. 3.1 the model was fitted to a variety of H I spectra from the THINGS survey (Walter et al. 2008). These galaxies were observed at high velocity resolution and under very low-noise conditions, and exhibited negligible baselines.

The fits are nearly always excellent. Local densities and voids in the H I distribution give rise to noisiness in the profiles, but residuals from these wiggles typically have amplitudes less than about 10% of the profile height. Such local deviations seem to have little effect on the bulk fitted properties.

Particularly important to note is the close fit to the slopes of the lines, which means that linewidth parameters of the profile such as W_{50} or W_{20} are accurately (within 2%) reproduced by

the best-fit model. There is a great deal of existing lore on how to correct such linewidth measurements so as to obtain estimates of the maximum rotation speed of the galaxy, which is the quantity desired for Tully-Fisher calculations. The availability of a good model proxy for the linewidth allows one to make use of such formulae without radical modification, even in cases of low spectral resolution or signal-to-noise ratio (S/N).

These correction formulae include a correction for instrumental broadening. Strictly speaking, this is no longer necessary if the linewidth is taken from the model profile, since this profile is what would actually be seen with no instrumental broadening (i.e., infinite resolution). Trials in Sect. 3.2 in which the model is fitted to spectra with coarser velocity binning indeed show no significant 'instrumental' bias in linewidth in the high-S/N limit. Biases do begin to show themselves as S/N decreases toward the limit of detectability, but they are always significantly less than those observed with linewidths measured directly from the noisy, coarsely-resolved profiles. As described in Sect. 3.2.7, correct Bayesian treatment of ensemble observations can circumvent ill effects from such low-S/N biases.

Comparison of the total flux parameter of the fitted model to that obtained by summing the observed flux densities indicates again that the model value of flux is an accurate proxy to the real one (within 1%). For about half the 34 galaxies fitted, the difference was less than the measurement uncertainty, which was nearly always significantly less than 1%. There is however a significant minority of galaxies for which the total flux is slightly underestimated. It is thought likely that this is because of the frequent occurrence of extended tails in the profile wings, which are underfitted by the Gaussian wings of the model.

Systemic velocities are returned within measurement errors and without bias by the model fits, even for asymmetric profiles.

The use of the remaining three model parameters (turbulent broadening Δv_{rand} , 'solid rotating' fraction f and asymmetry α) has been little explored. The Δv_{rand} value fitted is typically about 12 km s⁻¹ (see appendix B.2), which is 2 or 3 km s⁻¹ larger than typical values seen in spatially resolved maps of velocity distribution. As explained in Sect. 3.1.3, this is likely an artefact of the simplistic rotation curve implied by the model. Deviations of the rotation curves of real galaxies from this simple framework are readily absorbed into the Δv_{rand} parameter, tending therefore to inflate its value.

It might be of interest to see whether the f and α parameters can be useful as proxies for features of galaxy morphology, in analogy with studies such as those of Richter & Sancisi (1994) and Andersen & Bershadsky (2009), but this question is beyond the scope of the present paper.

In Sect. 3.2 it was shown that the model remains a useful way to extract parameters of interest even when the spectral resolution and S/N more nearly approximate those of the bulk of detected sources in surveys such as ALFALFA.

The final test performed on the model was to fit to a representative H I profile from the EBHIS survey (Winkel et al. 2010; Kerp et al. 2011). In contrast to the THINGS observations, this survey was done using a single dish antenna, and thus comes with a significant amount of baseline. This was dealt with in a Bayesian fashion by using the non-source areas of the relevant data cube to constrain the baseline priors. The success of this procedure is shown by the accurate agreement between the resulting fitted value of total flux and previous careful measurements for this galaxy (DDO 154).

Fitting the profile to individual pixels instead of to a spatial sum across several pixels allowed something of the spatial distribution of neutral hydrogen in this galaxy to be mapped in a low-

noise fashion, despite the intrinsically poor spatial resolution of the survey. The orientation of the galaxy's rotation axis could also be approximately determined via mapping the fitted value of the asymmetry parameter. The success of these fits shows that the model is not only suitable for H I profiles which result from a complete spatial integration across the extent of the galaxy, but also for partially integrated profiles.

Acknowledgements. IMS wishes to thank F Walter and J Kerp for generous provisions of non-public data, and is grateful to L Flöer, A Schröder and B Winkel for help and advice. We are grateful to the anonymous referee for helpful and informative comments.

References

- Andersen, D. R. & Bershad, M. R. 2009, *ApJ*, 700, 1626
- Barnes, D. G., Staveley-Smith, L., de Blok, W. J. G., et al. 2001, *MNRAS*, 322, 486
- Bhanot, G. 1988, *Rep. Prog. Phys.*, 51, 429
- Bicay, M. D. & Giovanelli, R. 1986, *AJ*, 91, 705
- Braun, R. 1997, *Ap. J.*, 484, 637
- Brewer, B. J., Marshall, P. J., Auger, M. W., et al. 2014, *MNRAS*, 437, 1950
- Carignan, C. & Purton, C. 1998, *Ap. J.*, 506, 125
- Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, *AJ*, 115, 1693
- D'Agostini, G. 2003, *Rep. Prog. Phys.*, 66, 1383
- de Blok, W. J. G., Walter, F., Brinks, E., et al. 2008, *AJ*, 136, 2648
- Duffy, A. R., Meyer, M. J., Staveley-Smith, L., et al. 2012, *MNRAS*, 426, 3385
- Fernández, X., van Gorkom, J. H., Hess, K. M., et al. 2013, *ApJ*, 770, L29
- Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2005, *AJ*, 130, 2598
- Gregory, P. C. & Lored, T. J. 1992, *ApJ*, 398, 146
- Hanson, K. M. & Cunningham, G. S. 1998, *Proc. SPIE*, 3338, 371
- Haynes, M. P. & Giovanelli, R. 1984, *AJ*, 89, 758
- Haynes, M. P., Giovanelli, R., Martin, A. M., et al. 2011, *AJ*, 142, 170
- Holwerda, B. W., Blyth, S.-L., & Baker, A. J. 2012, *Proc. IAU*, 284, 496
- Ianjamasimanana, R., deBlok, W., Walter, F., & Heald, G. H. 2012, *AJ*, 144, 96
- Johnston, S., Taylor, R., Bailes, M., et al. 2008, *Exp. Astr.*, 22, 151
- Jonas, J. L. 2009, *Proc. IEEE*, 97, 1522
- Kerp, J., Winkel, B., Ben Bekhti, N., Flöer, L., & Kalberla, P. M. W. 2011, *Astron. Nach.*, 332, 637
- Koribalski, B. & Staveley-Smith, B. 2009, *Survey Science Proposal*, Tech. rep., ASKAP
- Lewis, B. M. 1983, *AJ*, 88, 962
- Lored, T. J. 2012, *arXiv:1208.3036v2 [astro-ph.IM]*
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, *J. Chem. Phys.*, 21, 1087
- Meyer, M. 2009, in *Proceedings of Panoramic Radio Astronomy: Wide-field 1-2 GHz research on galaxy evolution*, ed. G. Heald & P. Serra
- Nelder, J. A. & Mead, R. 1965, *Comp. J.*, 7, 308
- Obreschkow, D., Croton, D., De Lucia, G., Khochfar, S., & Rawlings, S. 2009a, *ApJ*, 698, 1467
- Obreschkow, D., Klöckner, H.-R., Heywood, I., Levrier, F., & Rawlings, S. 2009b, *ApJ*, 706, 1890
- Penrose, R. 1955, *Proc. Camb. Phil. Soc.*, 51, 406
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P., eds. 1992, *Numerical Recipes in Fortran 77. The Art of Scientific Computing*, 2nd edn. (Cambridge, UK: Cambridge University Press)
- Richter, O.-G. & Sancisi, R. 1994, *A&A*, 290, L9
- Roberts, M. S. 1978, *AJ*, 83, 1026
- Saintonge, A. 2007, *AJ*, 133, 2087
- Springob, C. M., Haynes, M. P., Giovanelli, R., & Kent, B. R. 2005, *ApJS*, 160, 149
- Swaters, R. A., Schoenmakers, R. H. M., Sancisi, R., & van Albada, T. S. 1999, *MNRAS*, 304, 330
- Tift, W. G. & Cocke, W. J. 1988, *Ap. J. Suppl.*, 67, 1
- Tully, R. B. & Fisher, J. R. 1977, *A&A*, 54, 661
- Walter, F., Brinks, E., de Blok, W. J. G., et al. 2008, *AJ*, 136, 2563
- Willis, A. G. & Bregman, J. D. 1993, in *The Westerbork Synthesis Radio Telescope user documentation*, ed. O. M. Kolkman (Netherlands Foundation for Research in Astronomy), II-8-1 to II-8-13
- Winkel, B., Kalberla, P. M. W., Kerp, J., & Flöer, L. 2010, *ApJS*, 188, 488

Appendix A: The model in Fourier space

Appendix A.1: The transform components

As mentioned in Sect. 2.3, the model is generated in Fourier space partly to avoid singularities, partly to facilitate convolution, and partly to mimic the process by which real radio spectra are generated in an XF-type correlator. To follow the latter process requires the accurate construction of a synthetic discrete autocorrelation function Z_j for $j = -N$ to $N - 1$. This is done by sampling the Fourier transform $Z(\tau)$ of the profile model at appropriate intervals in τ , which is proportional to the autocorrelation lag time. The correct autocorrelation values result, provided only that the spectral line to be modelled has no significant power outside the chosen velocity interval $[v_{lo}, v_{hi}]$. This interval needs to be decided on before calculating the transforms. Together with the number of channels N , it defines the channel width Δv_{chan} .

The Fourier transform $Z(\tau)$ of the entire profile model is obtained by multiplying together the separate transforms $Z_{intrinsic}(\tau)$ and $G(\tau)$ of the expressions respectively for $s_{intrinsic}(v)$ and the dispersive convolver $g(v)$ given in equations 2 and 3. The transform of the intrinsic or undispersed profile evaluates to

$$Z_{intrinsic}(\tau) = 2S \exp(i\psi\tau) \left\{ \left[(1-f)J_0(\tau) + \frac{2f}{\tau}J_1(\tau) \right] + i\alpha \left[(1-f)J_1(\tau) + \frac{2f}{\tau} \left[\frac{2}{\tau}J_1(\tau) - J_0(\tau) \right] \right] \right\} \quad (A.1)$$

where J_0 and J_1 are respectively the zeroth and first order Bessel functions of the first kind, and

$$\psi = \frac{2(v_{hi} - v_{ctr})}{\Delta v}.$$

The transform of the dispersive-motion convolver gives

$$G(\tau) = \exp \left[-2 \left(\frac{\Delta v_{rand}}{\Delta v} \tau \right)^2 \right].$$

The correct sampling is at $\tau = j\Delta\tau$ for integer $j \in [-N, N - 1]$, where

$$\Delta\tau = \frac{\pi\Delta v}{2N\Delta v_{chan}}.$$

Recall that Δv in all these expressions is the linewidth parameter of the model, as described in Sect. 2.2.

Appendix A.2: Computing close to singularities

Note that we did not group all the J_0 and J_1 terms together in equation A.1. This is because terms with τ in the denominator need special treatment as $\tau \rightarrow 0$.

The expression $J_1(\tau)/\tau$ in equation A.1 is analogous to the sinc function. For values of τ close to zero it should be approximated by its power series

$$\frac{J_1(\tau)}{\tau} = \frac{1}{2} \left(1 - \frac{\tau^2}{8} + \frac{\tau^4}{192} - \dots \right).$$

Let us denote the last set of terms in equation A.1 as

$$E(\tau) = \frac{1}{\tau} \left[\frac{2}{\tau}J_1(\tau) - J_0(\tau) \right].$$

The components of this function also have singularities at $\tau = 0$ which cause computational difficulties as $\tau \rightarrow 0$. The power series for E evaluates to

$$E(\tau) = \frac{\tau}{8} \left(1 - \frac{\tau^2}{12} + \frac{\tau^4}{384} - \dots \right).$$

Appendix A.3: Derivatives

It can be convenient (for example when doing Levenberg-Marquardt optimization) to have expressions for the derivatives of the model with respect to each of the six parameters. We give the Fourier transforms of these here.

For convenience we define $C(\tau)$ such that the whole-model transform can be expressed as

$$Z(\tau) = Z_{\text{intrinsic}}(\tau) \times G(\tau) = 2S G(\tau) \exp(i\psi/\tau) C(\tau).$$

C thus contains all terms within the outer square bracket of equation A.1. The derivatives can then be given as

$$\begin{aligned} \frac{\partial Z}{\partial v_{\text{ctr}}} &= \frac{-2i\tau}{\Delta v} Z \\ \frac{\partial Z}{\partial \Delta v} &= 2S G \exp(i\psi/\tau) \frac{\partial C}{\partial \Delta v} \\ \frac{\partial Z}{\partial S} &= Z/S \\ \frac{\partial Z}{\partial \Delta v_{\text{rand}}} &= -4\Delta v_{\text{rand}} \left(\frac{\tau}{\Delta v} \right)^2 Z \\ \frac{\partial Z}{\partial f} &= 2S G \exp(i\psi/\tau) \frac{\partial C}{\partial f} \\ \frac{\partial Z}{\partial \alpha} &= 2S G \exp(i\psi/\tau) \frac{\partial C}{\partial \alpha}. \end{aligned}$$

The C derivatives are as follows:

$$\begin{aligned} \frac{\partial C}{\partial \Delta v} &= \frac{-\tau}{\Delta v} \left\{ (1-f)J_1(\tau) + 2fE(\tau) - \right. \\ &\quad \left. - i\alpha \left[(1-f) \left(J_0(\tau) - \frac{J_1(\tau)}{\tau} \right) + \frac{2f}{\tau} J_1(\tau) - \frac{6f}{\tau} E(\tau) \right] \right\} \\ \frac{\partial C}{\partial f} &= \left[-J_0(\tau) + \frac{2}{\tau} J_1(\tau) \right] + i\alpha [-J_1(\tau) + 2E(\tau)] \\ \frac{\partial C}{\partial \alpha} &= i[(1-f)J_1(\tau) + 2fE(\tau)]. \end{aligned}$$

Appendix B: Model fitting - technical issues

Appendix B.1: Priors in general

Bayes' theorem (equation 5) describes the posterior probability distribution of the model parameter values for a given set of data. This represents all that can be known about the parameters. However usually it suffices to settle for less than this. Typically one wishes to approximate the posterior via a single set of optimum parameter values coupled with widths of the posterior function in each dimension of the parameter space. This represents a full description of the posterior only if this function is an unskewed Gaussian; but it is often a good enough approximation. 'Optimum' can mean the values of the parameters at the maximum of the posterior, but mean values are often an acceptable proxy for these. Only if the posterior is very asymmetric will these two sets of values be significantly different.

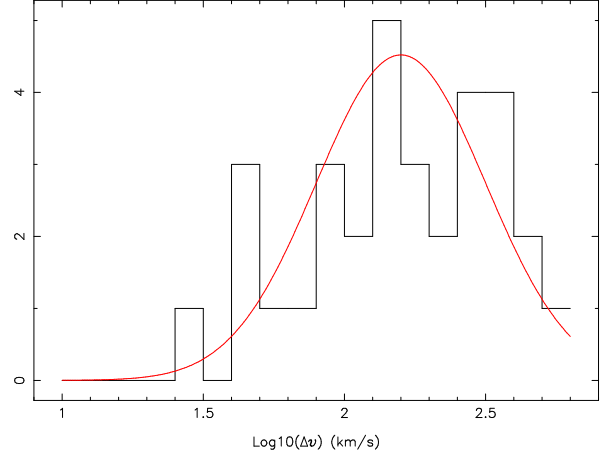


Fig. B.1. Distribution among the THINGS galaxies of Δv , the linewidth parameter of the fitted profiles. The red curve shows the prior which was chosen approximately to represent this distribution.

The posterior is a product of the likelihood and a function which represents prior knowledge of the parameter distribution. It is often easier to specify the likelihood than the priors. A typical example is where nothing definite at all is known about the likely distribution of the parameter values. What form of prior best encodes such a state of ignorance? There has been considerable debate about the correct forms for 'ignorance' priors (see for example references in D'Agostini 2003). The question depends however to some degree on what one wants to do with the posterior. If a value for the Bayesian evidence is what is needed, then priors have to be more carefully chosen: in particular, they must be integrable. Priors which are not integrable are named 'improper'. However in most cases discussed in the present paper, the location of the maximum in the posterior is all that is desired. If in addition the S/N of the data is high, the shape of the posterior will be dominated by a sharply-peaked likelihood, and the exact form of the priors becomes unimportant. One may even choose an improper prior without disadvantage.

Unbounded, thus improper, priors were used for most of the parameters in fitting the profile model to the original, high-S/N THINGS data. The fraction-solid and asymmetry parameters f and α have natural high and low bounds, but all the other parameters are unbounded on one or both sides.

Appendix B.2: Priors chosen for the low S/N spectra

The distributions of fitted parameters from the high-S/N fits were used to generate some priors for the low-S/N fits. Identical priors were applied in fitting both the semi-simulated spectra in Sect. 3.2 and the EBHIS spectra in Sect. 3.3.

The spectral-line fits performed on the 34 THINGS galaxies in Sect. 3.1 don't provide a very numerous sample, but some trends do become apparent. Using this information, non-flat priors were decided for four of the six model parameters, namely the linewidth Δv , the dispersion velocity width Δv_{rand} , the H I mass fraction in the core f , and the asymmetry parameter α . Histograms of the distribution of these four across the 34 THINGS fits are shown in Figs. B.1, B.2, B.3 and B.4.

A logarithmic histogram of the fitted linewidth parameter Δv is shown in Fig. B.1. Despite the single galaxy (DDO 53) with linewidth close to 10 km/s, a log-normal prior centred on 2.2 (i.e., at $10^{2.2}$ km/s) and of width 0.3 seemed a reasonable fit to the THINGS values.

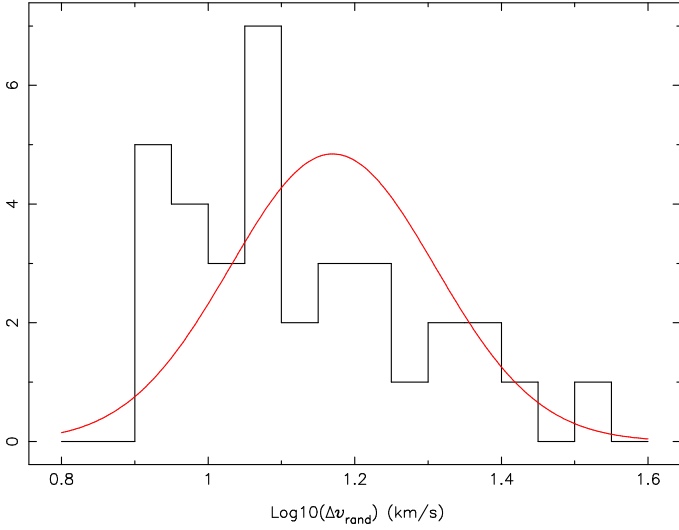


Fig. B.2. Distribution among the THINGS galaxies of Δv_{rand} , the velocity dispersion parameter of the fitted profiles. As in Fig. B.1, the red curve shows the prior.

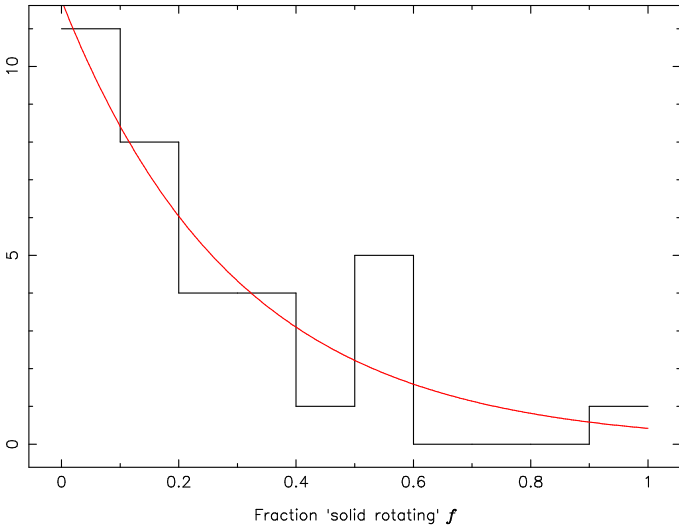


Fig. B.3. Distribution among the THINGS galaxies of f , the ‘fraction solid rotating’ parameter of the fitted profiles. As in Fig. B.1, the red curve shows the prior.

Figure B.2 shows the dispersion width. As already discussed, this is typically wider than would be expected from turbulence and thermal velocities alone. For this parameter we chose a log-normal prior function with a center in log space of 1.17 and a width of 0.14. It isn’t a particularly good fit to the histogram however: something with a sharper low-end cutoff and a longer tail would be better.

In Fig. B.3 a similar histogram for the core-fraction parameter is presented. The decrease in the incidence of this parameter as it increases in value is consistent with an exponentially decreasing prior, although a value of 0.3 was chosen for the characteristic constant so as to be consistent with the presence of the outlier in the histogram at the high end of the scale.

The final histogram in Fig. B.4 shows the asymmetry parameter. Values which depart from the midpoint by more than about 0.2 are seen to be unusual. The outliers here occur for spectra for which the model is not a good fit to the profile. A Gaussian with sigma equal to 0.2 was chosen for this prior.

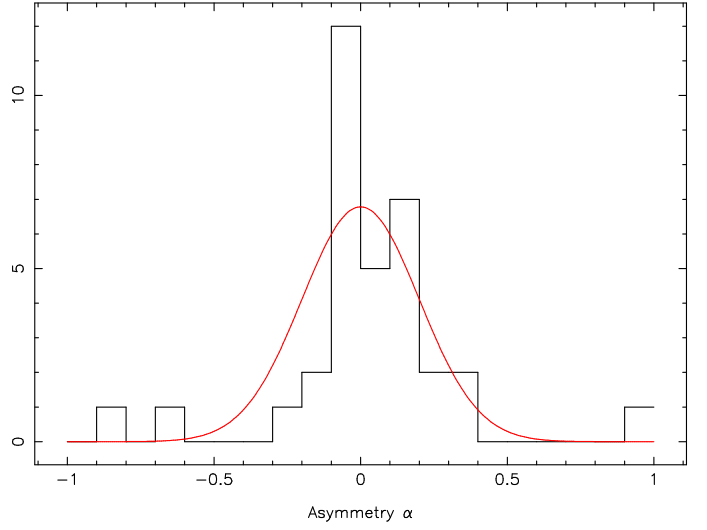


Fig. B.4. Distribution among the THINGS galaxies of α , the asymmetry parameter of the fitted profiles. As in Fig. B.1, the red curve shows the prior.

In addition to these four priors, and purely in order to exclude nonsense results due to fitting the profile to noise away from the true spectral line, we also chose a prior for the line centre parameter, this being a Gaussian centred on the true line centre (as determined from the high-S/N fits), and with a width somewhat arbitrarily set equal to 20 km s^{-1} . This is considered acceptable because we assume that in a real survey, the rough location of the line would already been found by other means. We remind the reader that we are not treating source detection here, but source characterisation.

For the remaining total-flux parameter, a flat prior was retained, with non-physical, negative-valued results excluded as before.

To summarize, the priors adopted in the present section were:

x	$p(x)$
v_{ctr}	$\exp(-0.5[(x - v_{\text{true}})/20]^2)/(20 \sqrt{2\pi})$
Δv	$\exp(-0.5[(\log 10(x) - 2.2)/0.3]^2)/(0.3 \ln 10 \sqrt{2\pi})$
S	$x > 0$
Δv_{rand}	$\exp(-0.5[(\log 10(x) - 1.17)/0.14]^2)/(0.14 \ln 10 \sqrt{2\pi})$
f	$0.3 \exp(-0.3x)$
α	$\exp(-0.5[x/0.2]^2)/(0.2 \sqrt{2\pi})$ for $ x \leq 1$, = 0 else.

Appendix B.3: Possible fitting problems

Appendix B.3.1: Degeneracy

The chosen set of parameters are convenient for many reasons but have the problem of not being orthogonal under all circumstances. For example, if the intrinsic linewidth Δv is small compared to the dispersion width Δv_{rand} , the f and α parameters begin to be poorly constrained. Figure B.5 serves to illustrate this point. And for at least one of the fits in Sect. 3.1, the best-fit value of α is close to unity, which causes nearly all the other parameters to be poorly constrained. These degeneracies can occur for both scenarios described in Sect. 2.4.1 and can cause convergence problems when trying to fit the profile. An MCMC is again the preferred way to deal with such cases.

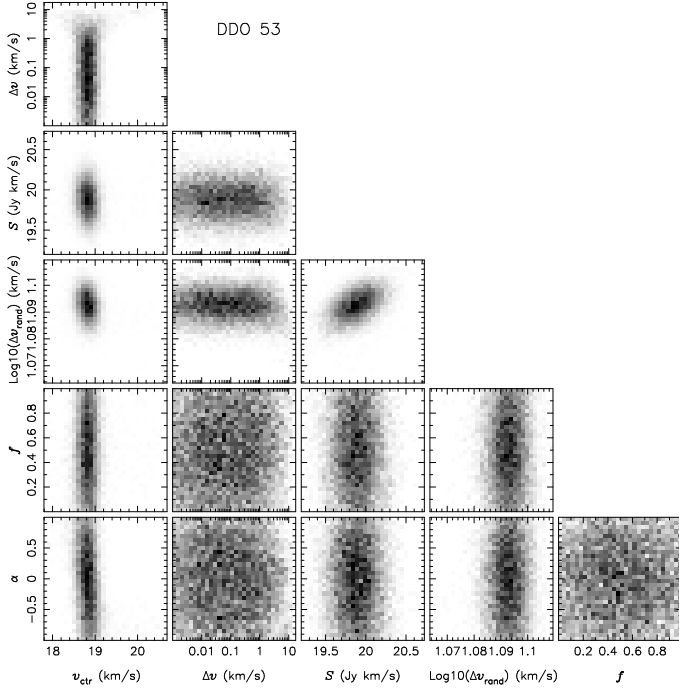


Fig. B.5. The shape of the posterior distribution for the six-parameter line profile model applied to DDO 53, as explored by the Markov-chain Monte Carlo (MCMC) procedure. Each pixel is shaded to represent the density of points in the MCMC.

Appendix B.3.2: Lag-space problems

Filtering in lag space in general causes the noise values in adjacent velocity channels to be correlated. The only common filtering scenario which avoids this is Hanning filtering followed by discarding every second channel. This was the scheme employed by Walter et al. (2008), so we have not had to address this problem. Other data sets are not so favourable though: for example, HIPASS made use of a 25% Tukey filter (Barnes et al. 2001) which introduces correlations between adjacent channels. The way to deal with this is called generalized least squares and involves reformulating the expression for χ^2 to have the form

$$\chi^2 = (\mathbf{s} - \mathbf{m})^T \mathbf{C}^{-1} (\mathbf{s} - \mathbf{m}). \quad (\text{B.1})$$

Here the T superscript indicates transpose and \mathbf{C} is the covariance matrix of the data channels. For uncorrelated data \mathbf{C} is of course diagonal and equation B.1 reverts to the standard expression. Equation B.1 seems straightforward but in fact \mathbf{C} is often singular for correlated data. This is effectively because there are now fewer than N degrees of freedom. Hanning smoothing for example reduces the number of degrees of freedom by a factor of 2, which is why one usually discards half the channels afterwards - they don't contain any new information. In general the solution to this is not to try to invert \mathbf{C} but to calculate instead its pseudo-inverse (Penrose 1955), which reduces the rank of the problem to the proper number. But this is not further elaborated here.

An alternative approach to the case of correlated noise would be to back-transform the data and perform the fitting in lag space, where there is no correlation between lag samples. This is a little complicated if the velocity channels have different standard deviations, which is the case in Sect. 3.1, so it has not been attempted here.

When fitting to a real spectral line, in principle both the velocity range and the number of channels should match that

in the original observation. This is because the end points where the autocorrelation function is truncated (particularly the low-frequency, i.e. high-velocity end) generate some ringing at places in the spectrum where its slope is steep (see e.g. Willis & Bregman 1993). If there is such ringing in the data, it's obviously advantageous if the model reproduces this artefact.

There were two reasons why, for many of the THINGS galaxies, this was difficult to do. Firstly, the cubes published by Walter et al. (2008) don't include all the observed channels. The starting channels had to be estimated from the tabulated description of the observations. Secondly, for some of the targets, Walter et al spliced together several observations taken with different velocity ranges. These cannot be accurately represented via a single simulated autocorrelation. These inaccuracies were however not regarded as significant.

Appendix B.4: Markov-chain Monte Carlo

A Markov-chain Monte Carlo (MCMC) is a sequence of steps within the parameter space. According to the Metropolis-Hastings algorithm (Metropolis et al. 1953), at each step a random displacement from the last good position is generated, and a test is applied to the value of the posterior at the resulting trial position. If the test is passed, the trial position is accepted as the next position in the chain; if not, a new trial is generated. It can be shown that the distribution of accepted positions will in time converge to that of the posterior.

This procedure is simple and robust to apply, but practical issues can arise, principally two: the speed of convergence, and how to know when the MCMC has converged. If the MCMC starts a long way from the centre of the posterior, and if the dimensions of the step size distribution are very different from those of the posterior, the rate of acceptances can be very low. This is the reason for using a LM fit in Sect. 3.1 to find the location and dimensions of the posterior maximum before starting the MCMC. After starting the MCMC, convergence speed was further improved via an iterative procedure, described as follows. The MCMC was divided into a sequence of separate chains. To begin with, the chains were short - only tens to hundreds of accepted steps were required. After finishing each chain, the covariance of the parameter points was calculated and used to specify the shape of an N -dimensional Gaussian step-size function for use in the next chain (N being the number of model parameters). The chains thus gradually converged at a faster rate, and more acceptances could be demanded. The final sequence of chains used had

$$N_A = \{30, 100, 200, 1000, 1000, 10000\},$$

where N_A is the number of acceptances required.

For a Gaussian posterior in a 6-dimensional parameter space, the fraction of accepted steps is expected to be a little less than 0.3 (estimated from Figs. 3 and 4 of Hanson & Cunningham 1998). This figure may be compared with the distribution of this fraction among the THINGS galaxies shown in Fig. B.6.

The parameter values and uncertainties listed in Table 1 were obtained by calculating the means and covariances of the final chain of 10^4 acceptances.

Others who employ the MCMC method commonly use far more points in the final chain than 10^4 ; typically also convergence tests are applied; sometimes multiple chains are used, with sparse sampling of the final chains, in order to avoid the obvious correlation between successive points. We have not felt any of this to be necessary. Convergence was judged by monitoring the

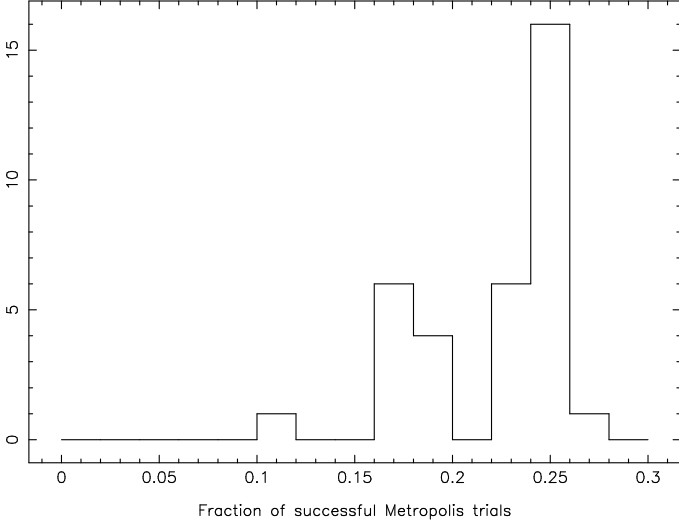


Fig. B.6. The distribution among the 34 THINGS galaxies of the fraction of successful trials returned by the Metropolis algorithm in the Markov-chain Monte Carlo profile-fitting procedure. The fraction was calculated from a chain of 10^4 successful trials after a burn-in chain of 2330 successes.

acceptance rate from chain to chain in the sequence, also by examining some plots of the successive parameter values. When no drift in the mean could be detected, it was judged that these chains have converged enough. The sequence given above was found to be adequate for all 34 profiles. Correlation was thought not to be a problem, since it is the ensemble of points which matters to us, and if there are enough points, then any sparse subset of them ought to have the same distribution as any other, and thus of the whole.

Appendix B.5: Calculation of the spectrum uncertainties

The THINGS spectra were formed by summing, for each velocity channel, the values of a subset of pixels in the corresponding plane of the data cube. The subset was chosen by Walter et al. (2008) to include only those pixels where some H I emission was detectable above the background. The uncertainty in each summed value can be expected to scale with the square root of the number of pixels summed, but calculating an exact value has some complications, because the noise values in adjacent spatial pixels are already correlated via convolution by the interferometer beam.

It is easiest to examine this in the Fourier domain. Given input uncorrelated noise of standard deviation σ which is then convolved by a beam $b_{j,k}$, the standard deviation or RMS σ_I of the output is given by

$$\begin{aligned}\sigma_I^2 &= \sigma^2 \langle \mathbf{B} \mathbf{B}^* \rangle \\ &= \sigma^2 \Sigma b^2 \text{ or } \sigma^2 \mathbf{b} \cdot \mathbf{b}\end{aligned}\quad (\text{B.2})$$

where \mathbf{B} is the discrete Fourier transform of the beam \mathbf{b} , $*$ signifies complex conjugation, and the angle brackets represent the mean value.

This is the state of the data in each image plane of the cube as we receive it. Note that in interferometry practice it is standard to normalize the beam such that $b_{0,0} = 1$. This allows one to read the flux density of point sources from the point image value $I_{j,k}$ at their locations, and gives rise to the expression ‘janskys per beam’ for the units of the image values. For an extended source,

the total flux density S can be approximated by

$$S \sim \frac{1}{A_{\text{beam}}} \sum_{j,k}^{A_{\text{patch}}} I_{j,k} \quad (\text{B.3})$$

where A_{patch} represents the subset of image pixels summed over, and A_{beam} is the ‘area’ in pixels of the beam, defined by

$$A_{\text{beam}} = \frac{1}{b_{0,0}} \sum b_{j,k}.$$

The problem we have is to calculate the uncertainty in the S values given by equation B.3, starting with a knowledge of \mathbf{b} and the value of σ_I . This becomes easy once we realize that equation B.3 can be thought of as a further convolution

$$\mathbf{S} = \frac{\mathbf{I} \star \mathbf{p}}{A_{\text{beam}}},$$

where the convolver \mathbf{p} equals unity for pixels within the patch or subset contributing to the sum, and zero otherwise. Bold font has been used to emphasize that all these quantities are images; the desired scalar value of spectrum flux density is in this notation given by $S_{0,0}$. The uncertainty in S can thus be related directly to the original σ by

$$\sigma_S^2 = \frac{\sigma^2}{A_{\text{beam}}^2} \langle (\mathbf{B} \mathbf{B}^*) (\mathbf{P} \mathbf{P}^*) \rangle,$$

where \mathbf{P} is the discrete Fourier transform of \mathbf{p} . More directly useful to us is to make use of equation B.2 to express σ_S in relation to the RMS σ_I of the cube planes:

$$\sigma_S^2 = \frac{\sigma_I^2}{A_{\text{beam}}^2} \frac{\langle (\mathbf{B} \mathbf{B}^*) (\mathbf{P} \mathbf{P}^*) \rangle}{\Sigma b^2}.$$

This can be calculated for any given \mathbf{p} , but it is of interest to look at its behaviour when $A_{\text{patch}} \gg A_{\text{beam}}$. In this case, we expect \mathbf{P} to be much more compact than \mathbf{B} , thus we can approximate

$$\begin{aligned}\sigma_S^2 &\sim \frac{\sigma_I^2}{A_{\text{beam}}^2} \frac{(B_{0,0} B_{0,0}^*) \langle \mathbf{P} \mathbf{P}^* \rangle}{\Sigma b^2} \\ &\sim \frac{\sigma_I^2}{A_{\text{beam}}^2} \frac{(\Sigma b)^2 A_{\text{patch}}}{\Sigma b^2}.\end{aligned}$$

For a Gaussian \mathbf{b} ,

$$\Sigma b^2 = \frac{b_{0,0}}{2} \Sigma b.$$

so we arrive at a final approximation

$$\sigma_S \sim \frac{\sigma_I}{b_{0,0}} \sqrt{\frac{2 A_{\text{patch}}}{A_{\text{beam}}}}.$$

Thus, with $b_{0,0}$ set to 1 according to usual interferometer practice, the ratio between the output and input uncertainties σ_S and σ_I is given, not by the square root of the ‘number of beams’ $A_{\text{patch}}/A_{\text{beam}}$, but by this quantity times root 2.

Appendix B.6: Deciding the order of the baseline model

In fitting to the EBHIS profiles we had to deal with spectra with non-flat baselines. For this purpose a sequence of Chebyshev polynomials T_j was added to the model, as described in Sect. 3.3:

$$s_{\text{base}}(v) = \sum_{j=1}^n a_j T_j(u)$$

where

$$u = 2 \left(\frac{v - v_{\text{lo}}}{v_{\text{hi}} - v_{\text{lo}}} \right) - 1.$$

One then has to decide at what value of n to truncate the sequence. The Bayesian methodology offers a way to decide this.

The posterior function $p(\mathbf{q}|\mathbf{y})$ giving the probability that the data \mathbf{y} is fitted by model parameter values \mathbf{q} is given in equations 5 and 7. So far we have not concerned ourselves with the evidence E , the denominator of equation 5, which is the integral over the parameter space of the numerator. If we change the number of parameters, however, the value of E will vary, and can be used to decide the optimum number.

In the present case the evidence integral has $6 + n$ dimensions, and considering for example Fig. B.5, we can see that the integrand, at least in the directions of the 6 profile-only parameters, is unlikely to be well behaved. However, the model is linear in the Chebyshev functions, which means that in these n directions, given Gaussian-distributed noise as we have, the likelihood function has also a Gaussian shape in these dimensions. In addition, the correlation between the n Chebyshev parameters and the 6 profile parameters may be expected to be small, particularly if we include rather more baseline than there are channels under the line profile. Therefore we can approximate the posterior probability function for the whole model by a product between the profile and baseline contributions:

$$p(\mathbf{q}|\mathbf{y}, n) = p(\mathbf{q}_{\text{profile}} + \mathbf{q}'|\mathbf{y}, n) \sim p(\mathbf{q}_{\text{profile}}|\mathbf{y}) \times [p(\mathbf{q}'|n) p(\mathbf{y}|\mathbf{q}', n)]. \quad (\text{B.4})$$

Here \mathbf{q}' represents the n baseline parameters, and the dependence of the probabilities on n has been made explicit.

Deciding the priors $p(\mathbf{q}'|n)$ presents its usual difficulties. Since the Chebyshev polynomials have a significant degree of orthogonality, we can approximate $p(\mathbf{q}'|n)$ by a product of n separate priors. We choose the same very broad prior for each Chebyshev term. ‘Broad’ of course implies slowly-varying compared to the likelihood function; thus each baseline prior may be replaced by its value at the posterior maximum, which we designate P_{base} . Since all n of them are separate and equal, $p(\mathbf{q}'|n) = P_{\text{base}}^n$.

The Gaussian likelihood for the n Chebyshev parameters has, following equation 7, the form

$$p(\mathbf{y}|\mathbf{q}', n) = p(\mathbf{y}|\mathbf{q}'_0, n) \exp \left(-[\mathbf{q}' - \mathbf{q}'_0]^T \frac{\mathbf{C}'^{-1}}{2} [\mathbf{q}' - \mathbf{q}'_0] \right). \quad (\text{B.5})$$

Here \mathbf{q}'_0 represents the values of the baseline parameters at the posterior maximum, and \mathbf{C}' is the covariance matrix of the posterior with respect to those parameters. The T superscript indicates transpose.

From equations 6, B.4 and B.5 the evidence evaluates to

$$E_n = \kappa_n P_{\text{base}}^n \sqrt{(2\pi)^n \det(\mathbf{C}')}.$$

where

$$\kappa_n = p(\mathbf{y}|\mathbf{q}'_0, n) \int d\mathbf{q}_{\text{profile}} p(\mathbf{q}_{\text{profile}}) p(\mathbf{y}|\mathbf{q}_{\text{profile}}).$$

Because of the weak coupling between the Chebyshev contribution and the profile model proper, we assume that κ_n can be decomposed into

$$\kappa_n \sim K p(\mathbf{q}_0|\mathbf{y}, n)$$

where K is a constant and $p(\mathbf{q}_0|\mathbf{y}, n)$ is the maximum value of the posterior, which can be measured at the end of the fit process. We cannot easily calculate K , but since it will be the same for any n , for purposes of comparing different n values we can simply evaluate

$$E'_n = \frac{E_n}{K} = p(\mathbf{q}_0|\mathbf{y}, n) P_{\text{base}}^n \sqrt{(2\pi)^n \det(\mathbf{C}')} \quad (\text{B.6})$$

Intuitively, as n increases, we would expect the fit to get better and better, reflected in an increased value of the posterior maximum, and thus of κ_n in equation B.6; but since P_{base} for a broad baseline prior is significantly less than 1, P_{base}^n will of course become geometrically smaller, and will eventually offset the rise in κ_n . This is what is known as the ‘Ockham’s Razor’ feature of the Bayesian methodology, in that it weights against a model with too many parameters (D’Agostini 2003; Gregory & Loredano 1992).

Shown in Fig. B.7 is the value of E'_n for different values of n . Two Gaussian priors for the Chebyshevs are tested, one having a width equal to 10 times the noise standard deviation (square points), the other being 10 times wider again (crosses). It can be seen that the width of the prior has small effect, provided only it is wide. In both cases the evidence shows a peak at $n = 6$. This was therefore the value chosen for the baseline fits.

Appendix B.7: Baseline priors for the EBHIS profiles

Many of the EBHIS cubes show significant baseline ripple. This must be fitted to obtain the most accurate estimates of the total flux (and to a lesser extent, the other parameters) of spectral lines. Of course the baseline variations continue smoothly through the spectral line itself, but if one includes the spectral line in an unconstrained baseline fit, one risks that the line itself will distort the baseline estimation. Traditionally this is accommodated by manually excluding the line channels from a separate baseline fit. Formally speaking however it is better to fit the same model to every channel, but to constrain the baseline parameters by priors. This has been our approach.

Gaussian priors were chosen for each of the Chebyshev amplitudes in the baseline model. The centre and spread for each were obtained in the following way. Firstly, the baseline model was fitted over the same channel range to the spectrum at every spatial pixel of the data cube, excluding only those pixels either where a significant source contribution was obvious, or at the edges of the cube where the noise was higher than average. Then, for each Chebyshev order, the centre of the prior was taken from the mean of the whole-cube fitted amplitudes, multiplied by the number of spatial pixels summed over the source at issue; and the spread (sigma value) of the prior was assigned from the standard deviation of these amplitudes, suitably corrected for the existing correlation between spatial pixels which is caused by the breadth of the telescope beam.

Appendix C: Graphs of the fits to all spectra

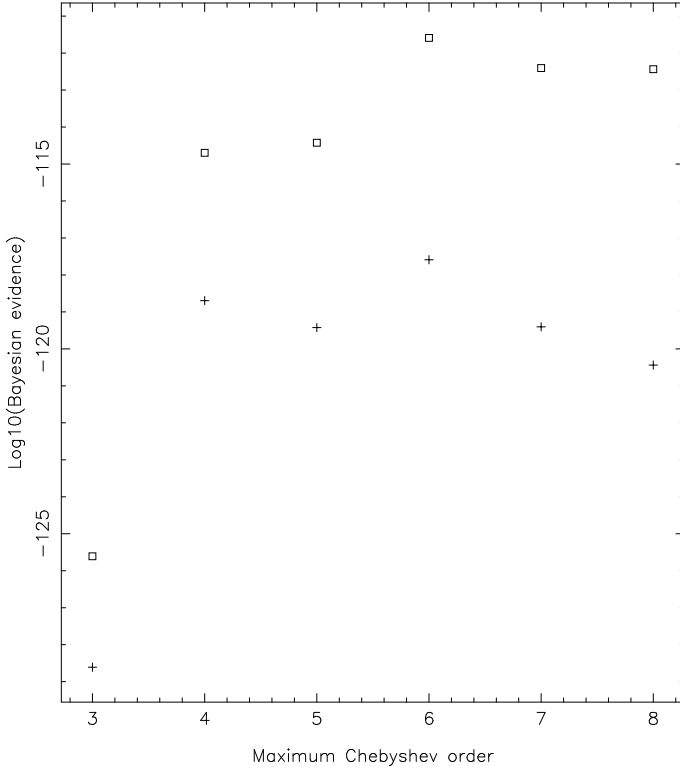


Fig. B.7. The Bayesian evidence for different maximum order of Chebyshev function. The fit was performed on an EBHIS spectrum for DDO 154.

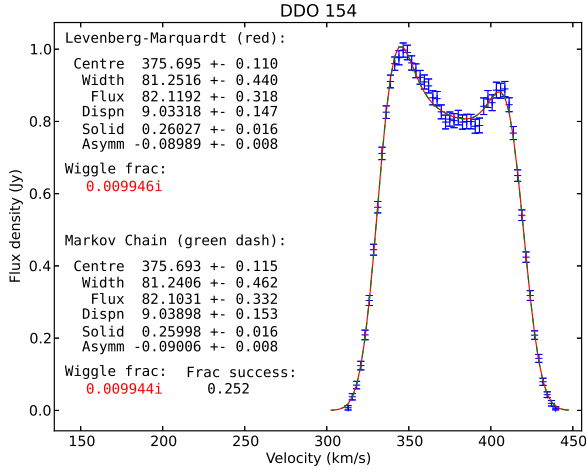


Fig. C.1.

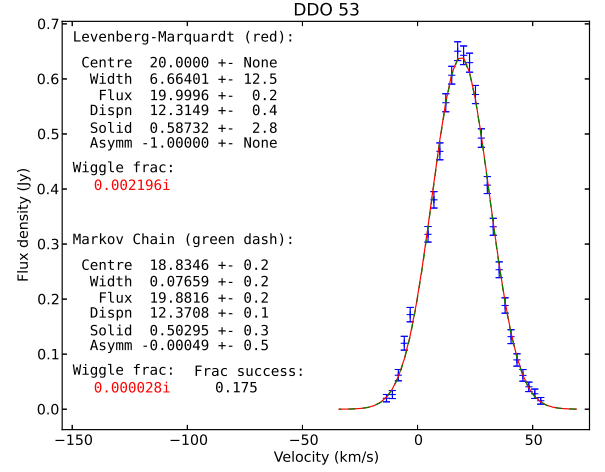


Fig. C.2.

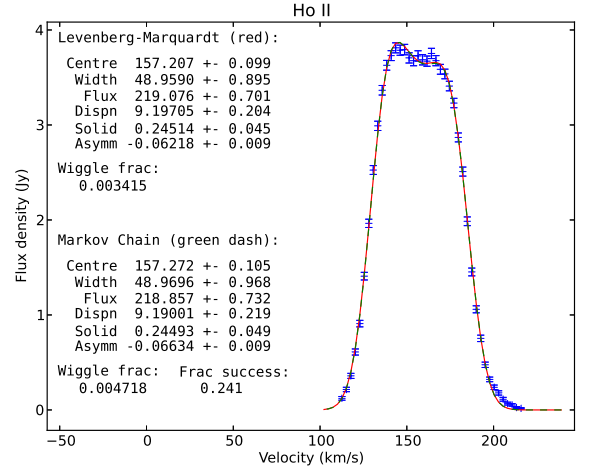


Fig. C.3.

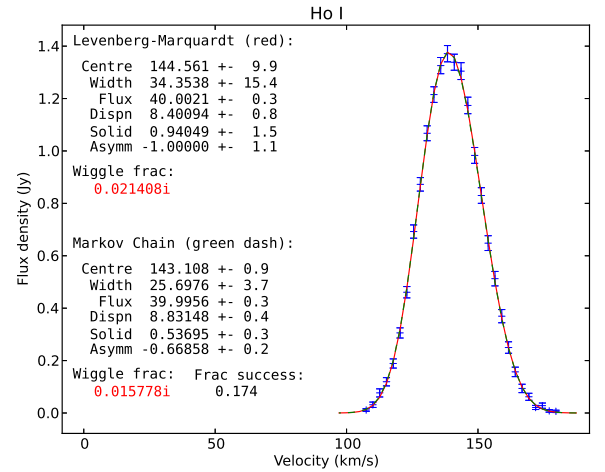


Fig. C.4.

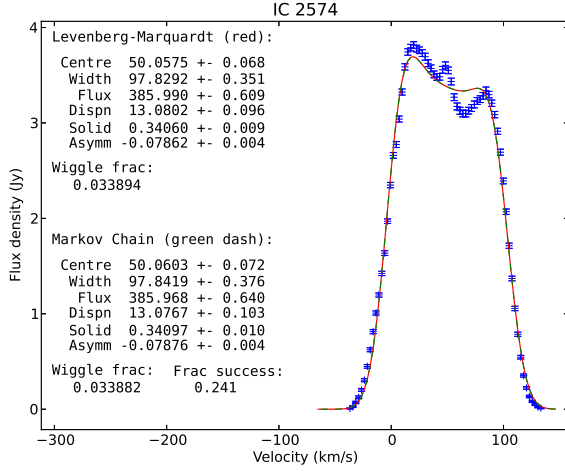


Fig. C.5.

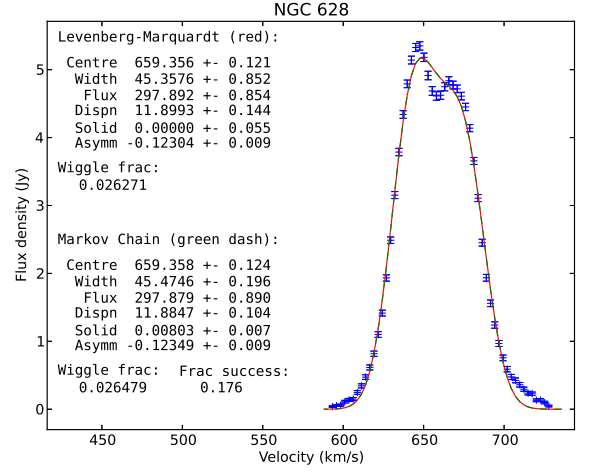


Fig. C.8.

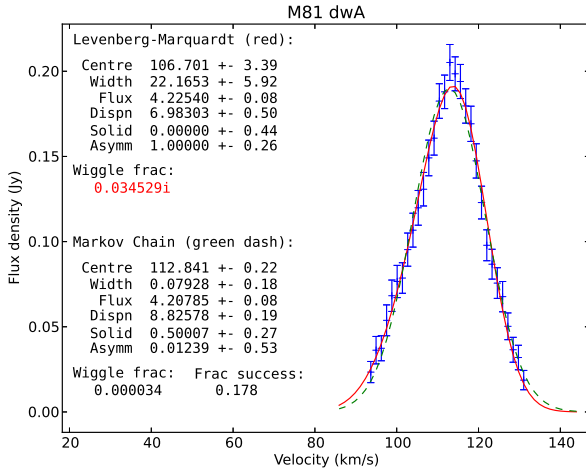


Fig. C.6.

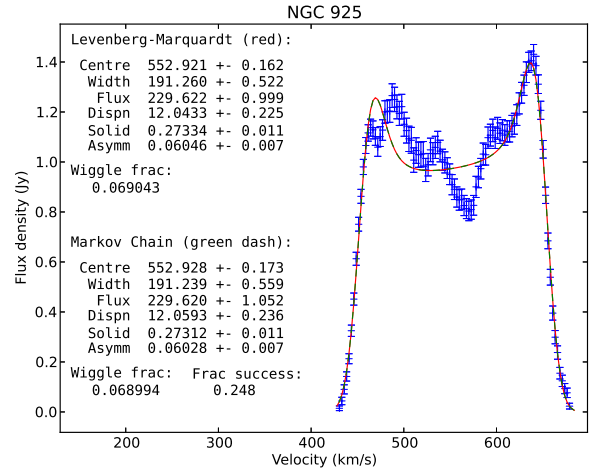


Fig. C.9.

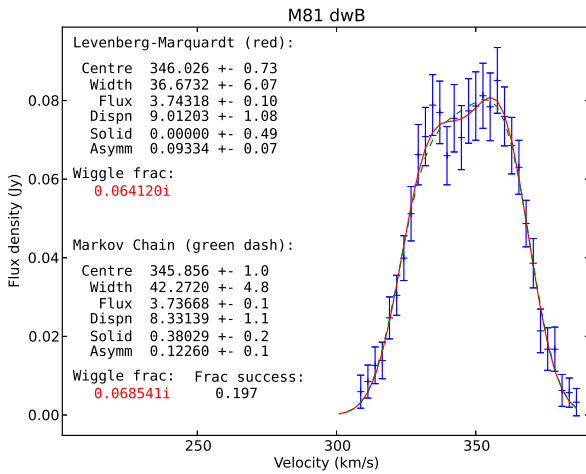


Fig. C.7.

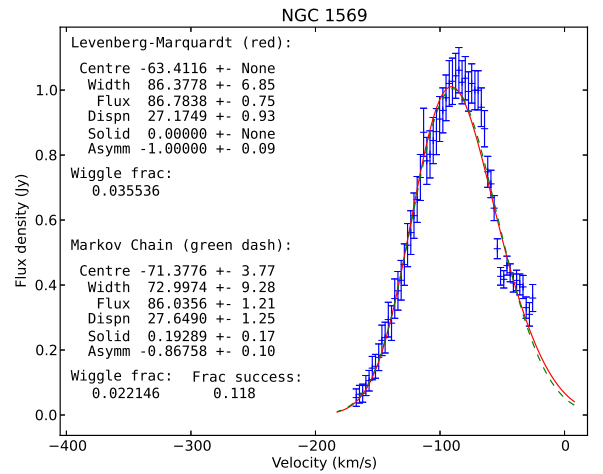


Fig. C.10.

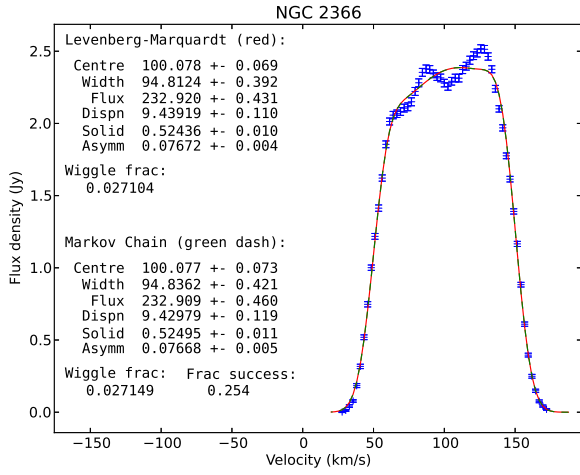


Fig. C.11.

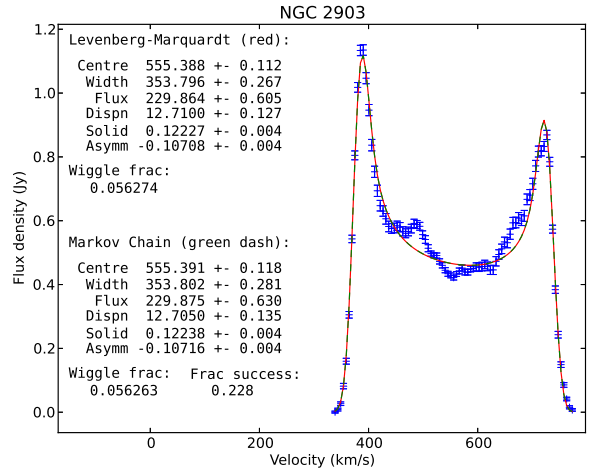


Fig. C.14.

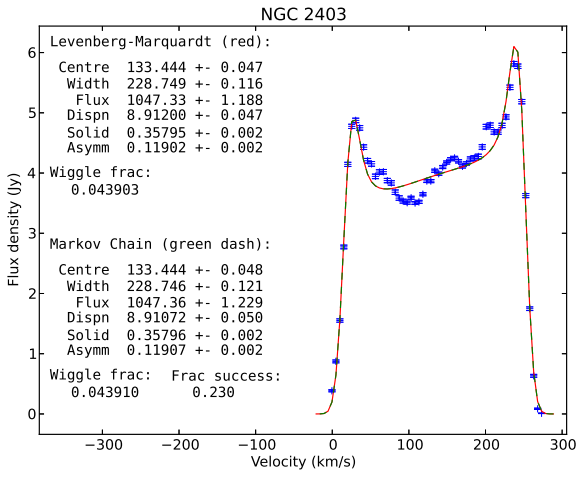


Fig. C.12.

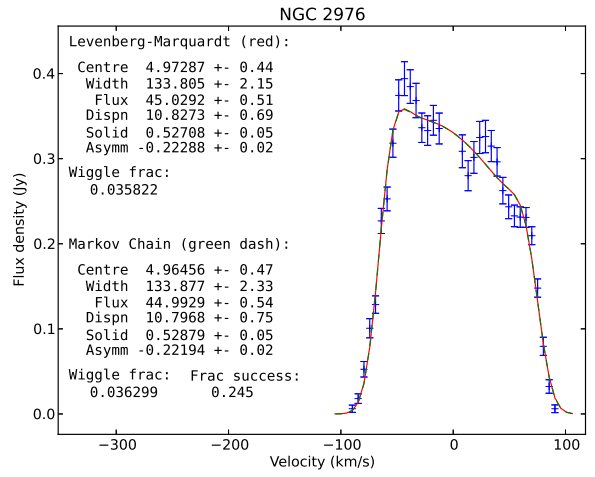


Fig. C.15.

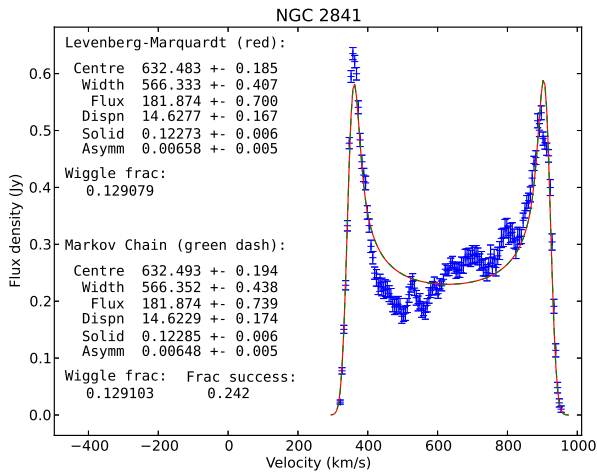


Fig. C.13.

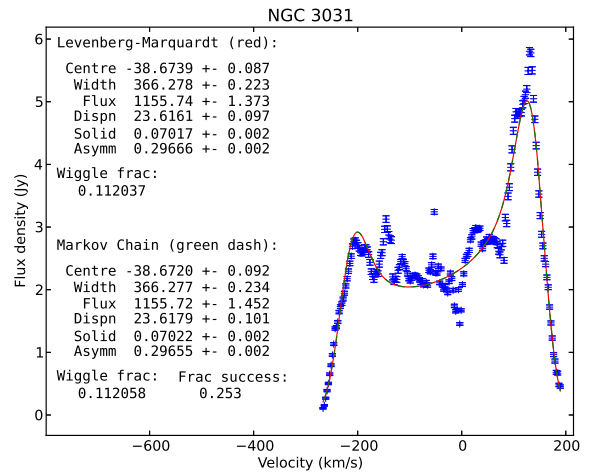


Fig. C.16.

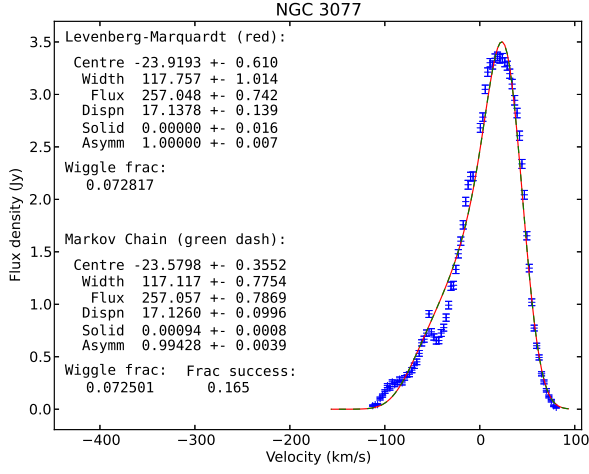


Fig. C.17.

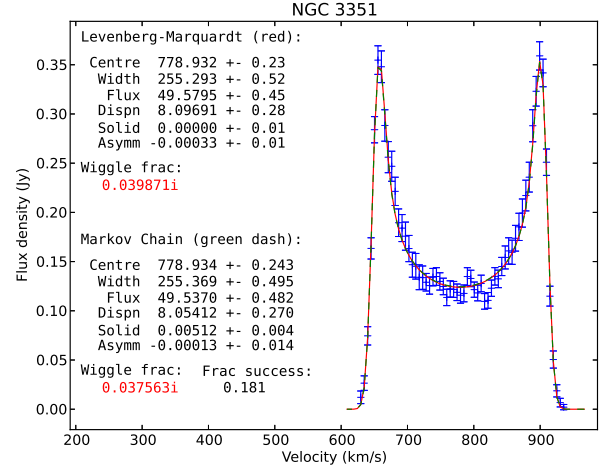


Fig. C.20.

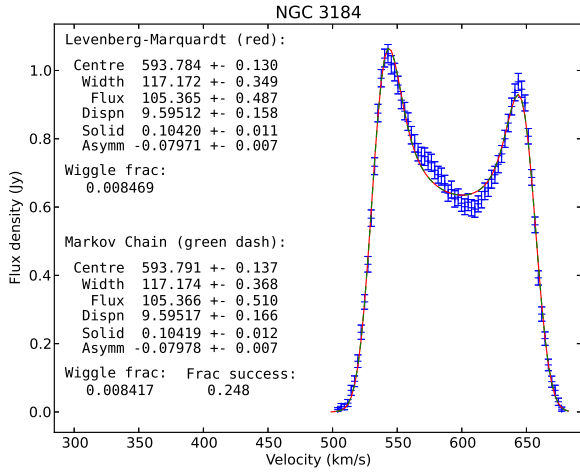


Fig. C.18.

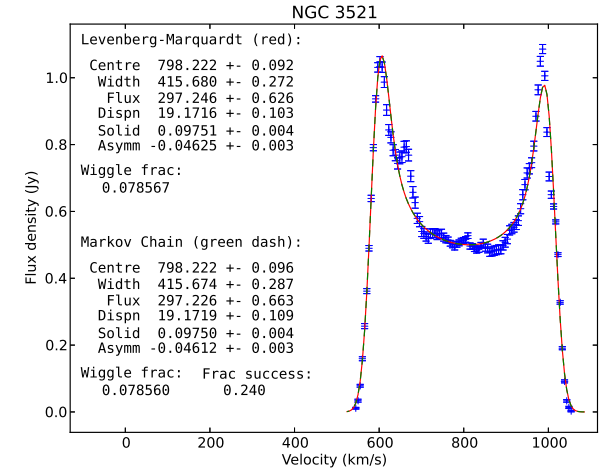


Fig. C.21.

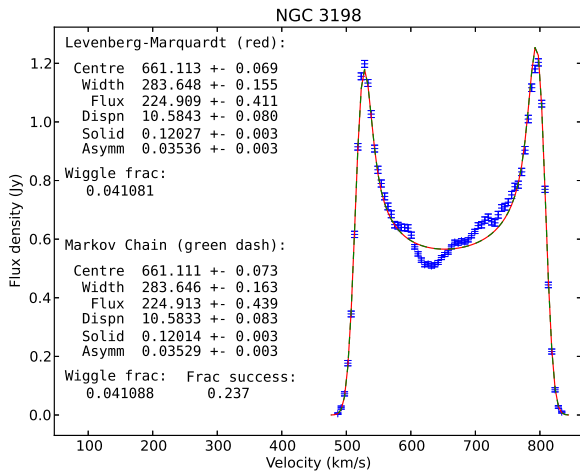


Fig. C.19.

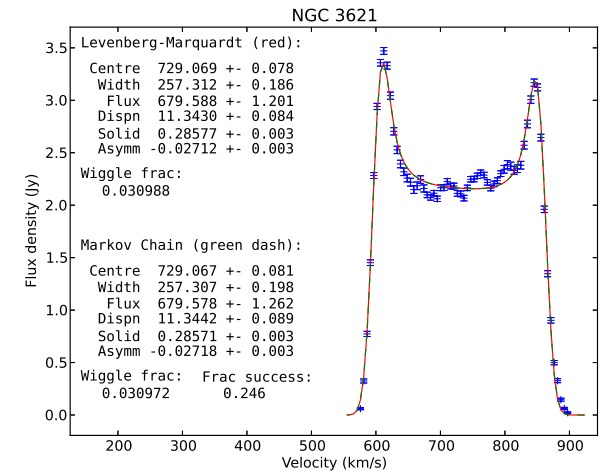


Fig. C.22.

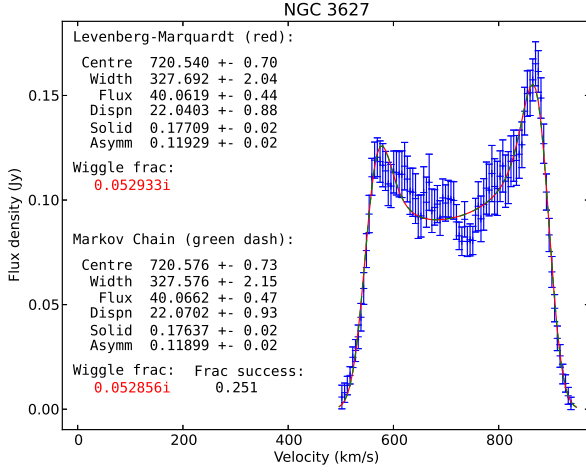


Fig. C.23.

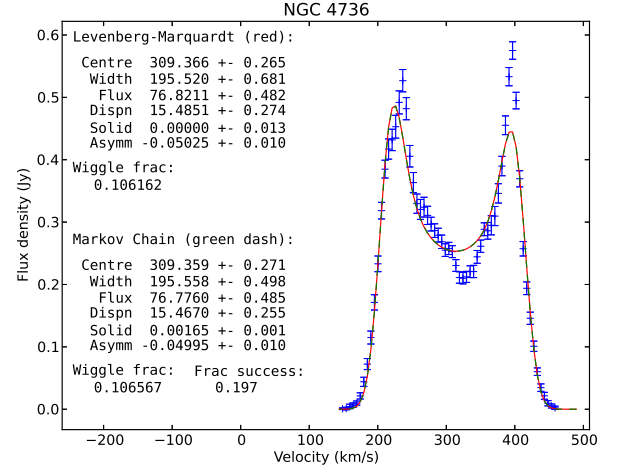


Fig. C.26.

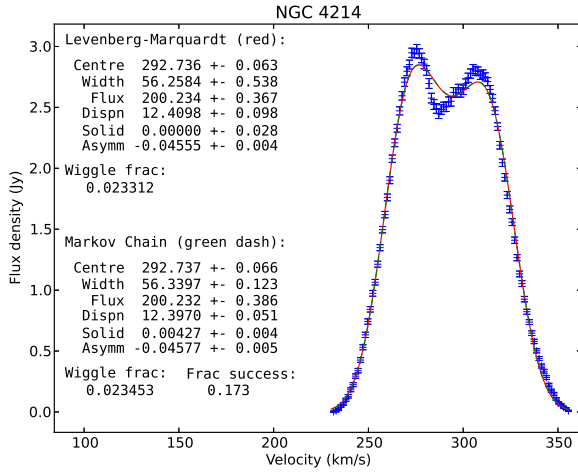


Fig. C.24.

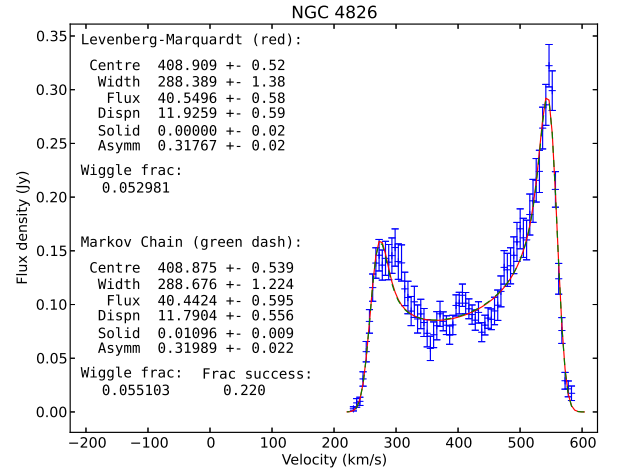


Fig. C.27.

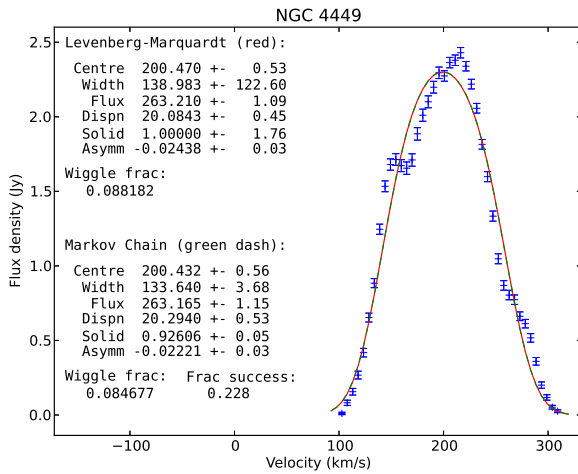


Fig. C.25.

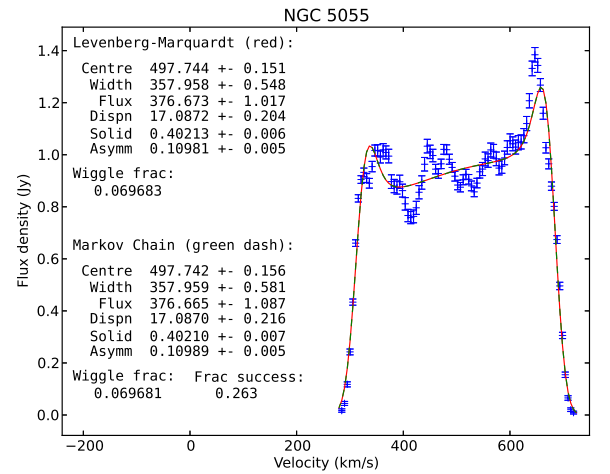


Fig. C.28.

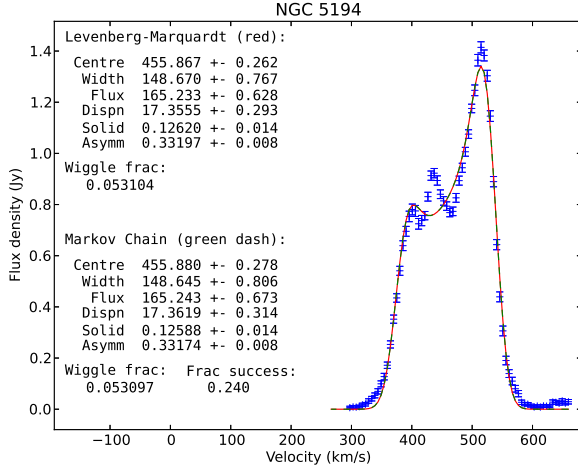


Fig. C.29.

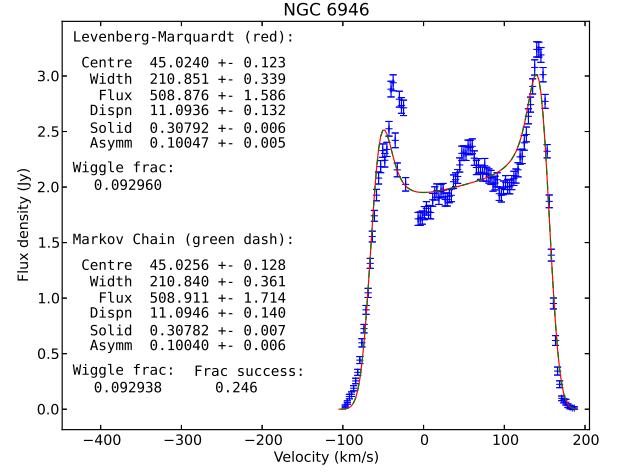


Fig. C.32.

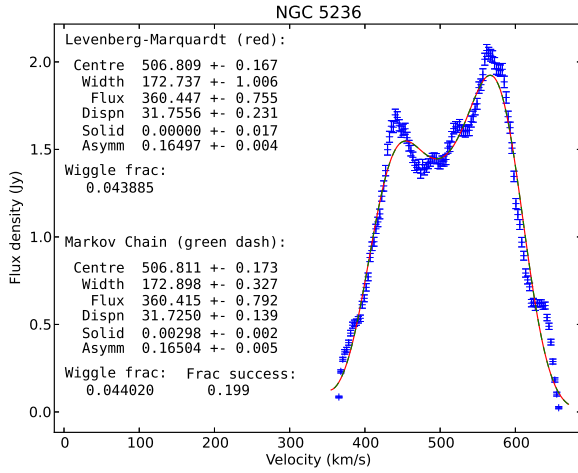


Fig. C.30.

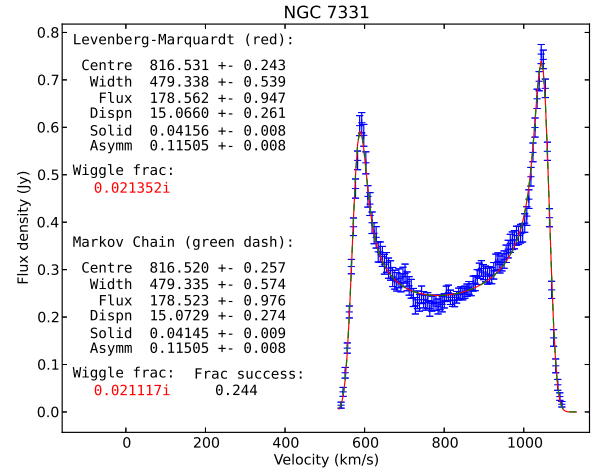


Fig. C.33.

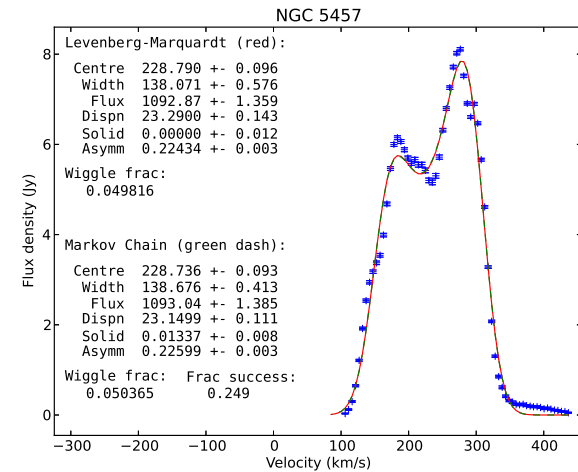


Fig. C.31.

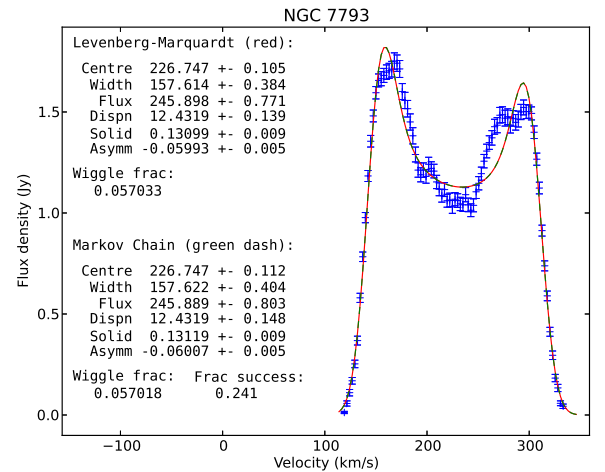


Fig. C.34.